

This Page Is Inserted by IFW Operations  
and is not a part of the Official Record

## **BEST AVAILABLE IMAGES**

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

**As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.**

**THIS PAGE BLANK (USPTO)**

Gaila Bunsul  
7

Access DB#

# SEARCH REQUEST FORM

Scientific and Technical Information Center

59224

Requester's Full Name: Dr. Bunsul, Gaila Examiner #: 72967 Date: 1/29/02  
Art Unit: 1142 Phone Number 303 395 Serial Number: 09160360  
Mail Box and Bldg/Room Location: CM/8A02 Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

\*\*\*\*\*  
Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc, if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: Molecular Complexes which modify the immune system

Inventors (please provide full names): Schaperke, Tomaltian

O'Hannon

Earliest Priority Filing Date: 3/28/06

\*For Sequence Searches Only\* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.

Please search

a composition comprising a cell wherein a  
molecular complex is bound to the surface of the cell,  
and the complex comprises  
4 fusion proteins.

two fusion proteins are <sup>Ig</sup>Heavy chain & a transmembrane  
polypeptide

and  
two fusion proteins are <sup>Ig</sup>light chain & a transmembrane  
polypeptide

2. Polypeptide sequence: Seq. 1: 1011/Gly-Glu-Leu-Gly-Gly Ser

Seq. 2: 1011/Gly-Glu-Leu-Gly-Gly Ser

Seq. 3: 1011/Gly-Gly-Gly-Thr-Ser Gly

## STAFF USE ONLY

Searcher: Point of Contact:

Alex Wadlaw

Searcher Phone #: Technical Info. Specialist

Searcher Extension: 2014 Tel: 303-4491

Date Searcher Picked Up: 7-1-02

Date Completed: 7-4-02

Searcher Prep & Review Time: \_\_\_\_\_

Clerical Prep Time: \_\_\_\_\_

Online Time: \_\_\_\_\_

## Type of Search

NA Sequence (#) \_\_\_\_\_

AA Sequence (#) 2

Structure (#) \_\_\_\_\_

Bibliographic /

Litigation \_\_\_\_\_

Fulltext \_\_\_\_\_

Patent Family \_\_\_\_\_

Other \_\_\_\_\_

## Vendors and cost where applicable

STN 1

Dialog \_\_\_\_\_

Questel/Orbit \_\_\_\_\_

Dr.Link \_\_\_\_\_

Lexis/Nexis \_\_\_\_\_

Sequence Systems \_\_\_\_\_

WWW/Internet \_\_\_\_\_

Other (specify) \_\_\_\_\_

**THIS PAGE BLANK (USPTO)**



Gencore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:00:16 : Search time 38.23 Seconds  
(without alignments)  
11.625 Million cell updates/sec

Title: US-09-642-660-10

Perfect score: 13 GCGC95.6

Sequence: 1 GCGC95.6

Scoring table: Gapop 10.0, Gapext 0.5

Searched: 52463 seqs, 74073290 residues

Total number of hits satisfying chosen parameters: 52463

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 100 summaries

Database: Geneseq1101\*

1:	/SID58/gcgdata/geneseq/AA1980.DMT*
2:	/SID58/gcgdata/geneseq/AA1981.DMT*
3:	/SID58/gcgdata/geneseq/AA1982.DMT*
4:	/SID58/gcgdata/geneseq/AA1983.DMT*
5:	/SID58/gcgdata/geneseq/AA1984.DMT*
6:	/SID58/gcgdata/geneseq/AA1985.DMT*
7:	/SID58/gcgdata/geneseq/AA1986.DMT*
8:	/SID58/gcgdata/geneseq/AA1987.DMT*
9:	/SID58/gcgdata/geneseq/AA1988.DMT*
10:	/SID58/gcgdata/geneseq/AA1989.DMT*
11:	/SID58/gcgdata/geneseq/AA1990.DMT*
12:	/SID58/gcgdata/geneseq/AA1991.DMT*
13:	/SID58/gcgdata/geneseq/AA1992.DMT*
14:	/SID58/gcgdata/geneseq/AA1993.DMT*
15:	/SID58/gcgdata/geneseq/AA1994.DMT*
16:	/SID58/gcgdata/geneseq/AA1995.DMT*
17:	/SID58/gcgdata/geneseq/AA1996.DMT*
18:	/SID58/gcgdata/geneseq/AA1997.DMT*
19:	/SID58/gcgdata/geneseq/AA1998.DMT*
20:	/SID58/gcgdata/geneseq/AA1999.DMT*
21:	/SID58/gcgdata/geneseq/AA2000.DMT*
22:	/SID58/gcgdata/geneseq/AA2001.DMT*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	33	100.0	12	20	AAW5467	LS170 polypeptide
2	33	100.0	18	20	AAW5464	LS170 polypeptide
3	33	100.0	19	20	AAW5471	LS170 polypeptide
4	33	100.0	61	22	AAW1383	Peptide #4303 enco
5	33	100.0	61	22	AAW1383	Peptide #4303 enco
6	33	100.0	93	22	AAW10376	Peptide #4403 enco
7	33	100.0	93	22	AAW10376	Peptide #4403 enco
8	33	100.0	256	20	AAW06408	Human secreted pro
9	33	100.0	256	20	AAW9443	LS170 polypeptide
10	33	100.0	256	21	AAW69164	Amino acid sequenc
11	33	100.0	256	22	AAW93721	Human polypeptide

12	33	100.0	256	22	AAW97366	Human LUNX protein
13	33	100.0	264	22	AAW41507	Human polypeptide
14	33	100.0	272	22	AAW41789	Human polypeptide
15	33	100.0	272	22	AAW41789	Human polypeptide
16	33	100.0	302	21	AAW11577	Arabidopsis thalia
17	33	100.0	302	21	AAW11575	Arabidopsis thalia
18	33	100.0	303	21	AAW11575	Arabidopsis thalia
19	33	100.0	348	22	AAW32858	Human heat shock p
20	33	100.0	348	22	AAW94509	Human protein sequ
21	33	100.0	378	19	AAW65428	Human secreted pro
22	33	100.0	499	21	AAW12967	Arabidopsis thalia
23	33	100.0	520	20	AAW93500	Human protein sequ
24	33	100.0	552	21	AAW12986	Human protein sequ
25	33	100.0	552	21	AAW12986	Human protein sequ
26	33	100.0	557	22	AAW94160	Arabidopsis thalia
27	33	100.0	582	22	AAW40004	Human protein sequ
28	33	100.0	602	22	AAW40004	Human protein sequ
29	33	100.0	625	22	AAW46032	S. epidemidis ope
30	33	100.0	653	22	AAW46032	Human polypeptide
31	33	100.0	2071	21	AAW84686	Human protein sequ
32	33	100.0	3025	22	AAW86196	Amino acid sequenc
33	33	100.0	3025	22	AAW90042	HIV-1 subtype C p
34	30	50.9	66	19	AAW7518	B. steerothromphi
35	30	50.9	81	21	AAW15636	Staphylococcus aur
36	30	50.9	83	21	AAW15636	Arabidopsis thalia
37	30	50.9	116	20	AAW27159	Arabidopsis thalia
38	30	50.9	116	20	AAW27159	Arabidopsis thalia
39	30	50.9	258	21	AAW94584	Zea mays protein f
40	30	50.9	258	21	AAW94584	Arabidopsis thalia
41	30	50.9	291	21	AAW31313	Human protein sequ
42	30	50.9	291	21	AAW31313	Arabidopsis thalia
43	30	50.9	315	21	AAW60616	Arabidopsis thalia
44	30	50.9	315	21	AAW60616	Arabidopsis thalia
45	30	50.9	356	22	AAW92510	Human protein sequ
46	30	50.9	372	22	AAW92109	C glutamicum prote
47	30	50.9	410	22	AAW81172	Mycobacterium tube
48	30	50.9	430	21	AAW54147	Arabidopsis thalia
49	30	50.9	430	21	AAW54147	Arabidopsis thalia
50	30	50.9	453	21	AAW61311	Arabidopsis thalia
51	30	50.9	453	21	AAW61311	Arabidopsis thalia
52	30	50.9	453	21	AAW61311	Arabidopsis thalia
53	30	50.9	453	21	AAW61311	Arabidopsis thalia
54	30	50.9	482	22	AAW84011	S. epidemidis ope
55	30	50.9	537	21	AAW8564	Shrimp white spot
56	30	50.9	538	18	AAW61300	Amino acid sequenc
57	30	50.9	538	18	AAW61300	Pseudomonas fluores
58	30	50.9	538	19	AAW69397	Pyricularia gene
59	30	50.9	538	19	AAW69397	Pyricularia gene
60	30	50.9	538	19	AAW69397	Pyricularia gene
61	30	50.9	538	19	AAW69397	Pyricularia gene
62	30	50.9	538	19	AAW69397	Pyricularia gene
63	30	50.9	538	19	AAW69397	Pyricularia gene
64	30	50.9	538	19	AAW69397	Pyricularia gene
65	30	50.9	538	19	AAW69397	Pyricularia gene
66	30	50.9	538	19	AAW69397	Pyricularia gene
67	30	50.9	538	19	AAW69397	Pyricularia gene
68	30	50.9	538	19	AAW69397	Pyricularia gene
69	30	50.9	538	19	AAW69397	Pyricularia gene
70	30	50.9	538	19	AAW69397	Pyricularia gene
71	30	50.9	538	19	AAW69397	Pyricularia gene
72	30	50.9	538	19	AAW69397	Pyricularia gene
73	30	50.9	538	19	AAW69397	Pyricularia gene
74	30	50.9	538	19	AAW69397	Pyricularia gene
75	30	50.9	538	19	AAW69397	Pyricularia gene
76	30	50.9	538	19	AAW69397	Pyricularia gene
77	30	50.9	538	19	AAW69397	Pyricularia gene
78	30	50.9	538	19	AAW69397	Pyricularia gene
79	30	50.9	538	19	AAW69397	Pyricularia gene
80	30	50.9	538	19	AAW69397	Pyricularia gene
81	30	50.9	538	19	AAW69397	Pyricularia gene
82	30	50.9	538	19	AAW69397	Pyricularia gene
83	30	50.9	538	19	AAW69397	Pyricularia gene
84	30	50.9	538	19	AAW69397	Pyricularia gene

ALIGNMENTS

85	29	87.9	38	22	AA884107	Immunomodulatory p
86	29	87.9	41	22	AA884107	Plant non-specific
87	29	87.9	45	20	AA119494	Human acid sequence
88	29	87.9	45	22	AA867214	Human acid sequence
89	29	87.9	45	22	AA867214	TF peptide #1 Syn
90	29	87.9	53	19	AA207990	Human neurofilament
91	29	87.9	55	20	AAV05524	80kE zipper Peptid
92	29	87.9	55	20	AAV05526	80kE zipper Peptid
93	29	87.9	55	20	AAV05528	80kE zipper Peptid
94	29	87.9	85	20	AAV05528	80kE zipper Peptid
95	29	87.9	91	14	AAV18023	Human acid sequence
96	29	87.9	93	22	AAV18220	Human acid sequence
97	29	87.9	93	22	AAV18220	Human acid sequence
98	29	87.9	98	17	AA892879	Peptide #5916 enco
99	29	87.9	100	21	AA892879	MN protein region
100	29	87.9	105	20	AAV13035	Arabidopsis thalis

## RESULT 1

AA95467 standard: peptide: 12 AA.

AA95467:

29-MAR-1999 (first entry)

LS170 polypeptide fragment.

LS170 gene: In vivo imaging: Lung disease: cancer: drug screening.

Homo sapiens.

WC9856951-A1.

17-DEC-1998.

11-JUN-1998: 98MO-US11601.

11-JUN-1997: 97US-0049183.

(ABB0) ABBOTT LAB.

Billings-medell PA, Cohen M, Colpitts TE, Friedman PN,

Gordon J, Granderson J, Hodges SC, Klags MR, Kratochwill JD,

Roberts-rapp L, Russell JC, Stroupe SD;

WPI: 1999-060335/05.

New LS170 nucleic acid from lung tissue - useful for detecting,

monitoring, preventing and treating lung disease, particularly

cancer

Claim 17: Page 98: 120pp: English.

Sequences AA95463-471 represent LS170 polypeptide fragments which contain at least on epitope in their sequences. The present invention relates to detection of a target LS170 polynucleotide that comprises has at least 50 percent identity with any of the specific nucleic acid that which are fragments derived from various clones of LS170 gene. The LS170 nucleic acid fragments represent a set of contiguous, partially overlapping sequences transcribed from lung tissue. They are used for diagnosis, staging, monitoring, in vivo imaging, prevention and treatment of lung disease, specifically cancer, and to indicate predisposition to antigens, or anti-LS170 antibodies. The LS170 polynucleotide, LS170 nucleic acid, or anti-LS170 antibodies are used to express recombinant polypeptides are used to raise Ab and for drug screening. LS170-related nucleic acid can be used to isolate related sequences; as standards and reagents in

CC assays; as targets for drug screening, and as components or targets for CC therapy, e.g. as antisense, ribozyme or triplex-forming agents. Ab can be used to deliver therapeutic agents to LS170-expressing cells; directly as therapeutic agents (by neutralizing LS170 polypeptides); in competitive binding drug screens, and to generate anti-idiotypic antibodies for use in rational drug design.

Sequence 12 AA:

Query Match 100.0%; Score 33; DB 20; Length 12;

Match Local Similarity 100.0%; Pred. No. 26;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GGGTSG 6

111111

DB 6 99GTSG 11

## RESULT 2

AA95464 standard: peptide: 18 AA.

AA95464:

29-MAR-1999 (first entry)

LS170 polypeptide fragment.

LS170 gene: In vivo imaging: Lung disease: cancer: drug screening.

Homo sapiens.

WC9856951-A1.

17-DEC-1998.

11-JUN-1998: 98MO-US11601.

11-JUN-1997: 97US-0049183.

(ABB0) ABBOTT LAB.

Billings-medell PA, Cohen M, Colpitts TE, Friedman PN,

Gordon J, Granderson J, Hodges SC, Klags MR, Kratochwill JD,

Roberts-rapp L, Russell JC, Stroupe SD;

WPI: 1999-060335/05.

New LS170 nucleic acid from lung tissue - useful for detecting,

monitoring, preventing and treating lung disease, particularly

cancer

Claim 17: Page 97: 120pp: English.

Sequences AA95463-471 represent LS170 polypeptide fragments which contain at least on epitope in their sequences. The present invention relates to detection of a target LS170 polynucleotide that comprises has at least 50 percent identity with any of the specific nucleic acid that which are fragments derived from various clones of LS170 gene. The LS170 nucleic acid fragments represent a set of contiguous, partially overlapping sequences transcribed from lung tissue. They are used for diagnosis, staging, monitoring, in vivo imaging, prevention and treatment of lung disease, specifically cancer, and to indicate predisposition to antigens, or anti-LS170 antibodies. The LS170 polynucleotide, LS170 nucleic acid, or anti-LS170 antibodies are used to express recombinant polypeptides are used to raise Ab and for drug screening. LS170-related nucleic acid can be used to isolate related sequences; as standards and reagents in assays; as targets for drug screening, and as components or targets for therapy; e.g. as antisense, ribozyme or triplex-forming agents. Ab can be

CC used to deliver therapeutic agents to LS170-expressing cells; directly  
 CC as therapeutic agents (by neutralising LS170 polypeptides); in  
 CC competitive binding drug screens, and to generate anti-idiotypic  
 CC antibodies for use in rational drug design.

XX Sequence 18 AA:

Query Match 100.0%; Score 33; DB 20; Length 18;  
 Best Local Similarity 100.0%; Pred. No. 37;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 |||||  
 DB 7 99GTSG 12

# RESULT 3

AA095471 standard; peptide; 19 AA.

XX AA095471:

XX 29-MAR-1999 (first entry)

XX LS170 polypeptide fragment.

XX LS170 gene; in vivo imaging; lung disease; cancer; drug screening.

XX Homo sapiens.

XX W09856951-A1.

XX 17-DEC-1998.

XX 11-JUN-1998; 98MO-US11601.

XX 11-JUN-1997; 97US-0049183.

XX (ABBO) ABBOTT LAB.

XX Billing-metel PA, Cohen M, Colpitta TL, Friedland PN;  
 PI Gordon J, Grandoson, Hodges SC, Klass MR, Kretschwill JD;  
 PI Roberts-Trapp L, Russell JC, Stroupe SD;

XX WPI; 1999-06035/05.

XX New LS170 nucleic acid from lung tissue - useful for detecting,  
 PT monitoring, preventing and treating lung disease, particularly  
 PT cancer

XX Claim 17; Page 98; 120pp; English.

XX Sequences AA095463-471 represent LS170 polypeptide fragments which  
 CC contain at least one epitope in their sequences. The present invention  
 CC relates to detection of a target LS170 polynucleotide that comprises  
 CC a test sample with at least one LS170-specific nucleic acid that  
 CC has at least 50 percent identity with any of the sequences (AA00801-09)  
 CC which are fragments derived from various clones of LS170 gene. The LS170  
 CC nucleic acid fragments represent a set of contiguous, partially  
 CC overlapping sequences. The present invention provides a method for  
 CC diagnosis, screening, monitoring, in vivo imaging, prevention and treatment  
 CC of lung disease, specifically cancer, and to indicate predisposition to  
 CC such disease. Particularly detection of LS170 polynucleotide, LS170  
 CC antigens, or anti-LS170 antibodies is indicative of disease. Cells  
 CC transformed with an expression system comprising the LS170 nucleic acid  
 CC sequences are used to express recombinant polypeptides. The polypeptides  
 CC are used as standards in assays for detecting LS170 polypeptides in  
 CC assays; as targets for drug screening, and as components or targets for  
 CC therapy, e.g. as antisense, ribozyme or triplex-forming agents. Ab can be  
 CC used to deliver therapeutic agents to LS170-expressing cells; directly  
 CC as therapeutic agents (by neutralising LS170 polypeptides); in

CC competitive binding drug screens, and to generate anti-idiotypic  
 CC antibodies for use in rational drug design.

XX Sequence 19 AA:

Query Match 100.0%; Score 33; DB 20; Length 19;  
 Best Local Similarity 100.0%; Pred. No. 39;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 |||||  
 DB 7 99GTSG 12

# RESULT 4

AA035315 standard; protein; 47 AA.

XX AA035315:

XX 17-OCT-2001 (first entry)

XX Peptide #7552 encoded by probe for measuring placental gene expression.

XX Probe; microarray; human; placenta; antenatal diagnosis;

XX genetic disorder.

XX Homo sapiens.

XX W020015727-A2.

XX 09-AUG-2001.

XX 30-JAN-2001; 2001MO-US00663.

XX 04-FEB-2000; 2000US-0180312.

XX 26-MAY-2000; 2000US-0204456.

XX 30-JUN-2000; 2000US-0608408.

XX 03-AUG-2000; 2000US-0632366.

XX 21-SEP-2000; 2000US-0234687.

XX 27-SEP-2000; 2000US-0235359.

XX 04-OCT-2000; 2000US-0024263.

XX (MOLE) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

XX WPI; 2001-488897/53.

XX Human genome-derived single exon nucleic acid probes useful for  
 PT analyzing gene expression in human placenta -

XX Claim 27; SEQ ID NO 33764; 554pp; English.

XX The present invention relates to single exon nucleic acid probes (SENP;  
 CC see A013315-A157546). The present sequence is a peptide encoded by one  
 CC such probe. The probes are useful for producing a microarray for  
 CC predicting, measuring and displaying gene expression in samples  
 CC from human placenta. The probes are useful for antenatal diagnosis of  
 CC human genetic disorders.

XX Sequence 47 AA:

Query Match 100.0%; Score 33; DB 22; Length 47;  
 Best Local Similarity 100.0%; Pred. No. 46;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 |||||  
 DB 26 99GTSG 31



PR 26-MAY-2000; 2000US-0207456.  
 PR 30-JUN-2000; 2000US-0608408.  
 PR 03-AUG-2000; 2000US-0632366.  
 PR 21-SEP-2000; 2000US-0234687.  
 PR 27-SEP-2000; 2000US-0234687.  
 PR 04-OCT-2000; 2000US-0024253.  
 XX  
 PA (MOLE-) MOLECULAR DYNAMICS INC.  
 XX  
 PI Penn S5, Hanzel DK, Chen W, Rank DR:  
 XX  
 DR WPI: 2001-476286/51.  
 XX  
 PT Novel single exon nucleic acid probe used to measuring gene expression  
 PT in a human breast -  
 XX  
 PS Claim 27; SEQ ID No 14258; 322pp; English.  
 XX  
 CC The present invention relates to novel single exon nucleic acid probes  
 CC (see AAI0010-AA110067). The present sequence is a peptide encoded by one  
 CC human breast sample, where the probe hybridises at high stringency to a  
 CC nucleic acid expressed in the human breast. The probes are useful for  
 CC predicting, diagnosing, grading, staging, monitoring and prognosing  
 CC diseases of the human breast, particularly those diseases with polygenic  
 CC aetiology. The diseases include: breast cancer, disorders of development,  
 CC inflammatory diseases of the breast, fibrocystic changes, proliferative  
 CC breast disease and non-carcinoma tumours.  
 CC The human breast sample used in the present invention did not form part of the printed  
 CC specification but was obtained in electronic format directly from WPIPO  
 CC at ftp.wipo.int/pub/published\_pct\_sequences.  
 CC  
 SQ Sequence 93 AA:  
 100.0%; Score 33; DB 22; Length 93;  
 Query Match Best Local Similarity 100.0%; Pred. No. 1; e-402;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 46 999tsg 51  
 100.0%; Score 33; DB 22; Length 93;  
 Query Match Best Local Similarity 100.0%; Pred. No. 1; e-402;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 46 999tsg 51  
 RESULT 8  
 AA06408 standard; Protein: 256 AA.  
 XX ID AA06408 standard; Protein: 256 AA.  
 XX AC AA06408;  
 XX PT 20-SEP-1999 (first entry)  
 XX DE Human secreted protein nh796\_1.  
 XX KW Secreted protein: nh796\_1; human; brain; thalamus; therapy;  
 XX diagnosis.  
 XX OS Homo sapiens.  
 XX SS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 7..19 /note="Predicted leader/signal sequence"  
 XX FT Protein 20..256 /note="mature protein"  
 XX FT W09935252-A2.  
 XX PD 15-JUL-1999.  
 XX PF 31-DEC-1998; 98WO-US27903.  
 XX PR 30-DEC-1998; 98US-0222653.  
 PR 02-JAN-1998; 98US-0070346.

XX (GENY ) GENETICS INST INC.  
 PA Agostino MJ, Collins-Racie LA, Jacobs K, Lavaille ER:  
 XX McCoy JM, Herzberg D, Steininger NJ, Treacy W:  
 XX WPI: 1999-419350/35.  
 DR N-PSDB: AAX59356.  
 XX  
 PT New polynucleotides encoding secreted human proteins  
 PT  
 PS Claim 25; Page 96-97; 100pp; English.  
 XX  
 CC This sequence represents a novel human secreted protein, termed  
 CC nh796\_1. The sequence is predicted from a full-length cDNA clone  
 CC (see AAX59356) isolated from a human adult brain (thalamus) cDNA  
 CC library. The invention provides cDNA clones (see AAX59352-58)  
 CC encoding novel secreted proteins (see AA06404-10) of the human  
 CC testis, brain and foetal kidney. The polynucleotides and  
 CC proteins are predicted to have biological functions. It would  
 CC be expected that the proteins encoded by the present sequence would  
 CC conditions in humans and animals, although no supporting data are  
 CC provided. Suggested activities include nutritional, cytokine, cell  
 CC proliferation or differentiation, immunostimulant (e.g. as  
 CC vaccine) or immunosuppressive, haematopoietic regulating, tissue  
 CC growth, activin or inhibin, chemotactic or chemokinetic,  
 CC haemostatic, thrombolytic, receptor/ligand, antiinflammatory,  
 CC gadarhin or tumour invasion suppressor, and tumour inhibition  
 CC activity.  
 CC  
 SQ Sequence 256 AA:  
 100.0%; Score 33; DB 20; Length 256;  
 Query Match Best Local Similarity 100.0%; Pred. No. 3; e-02;  
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 80 999tsg 85  
 100.0%; Score 33; DB 20; Length 256;  
 Query Match Best Local Similarity 100.0%; Pred. No. 3; e-02;  
 Matches 0; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 80 999tsg 85  
 RESULT 9  
 AA095463 standard; Protein: 256 AA.  
 XX ID AA095463 standard; Protein: 256 AA.  
 XX AC AA095463;  
 XX PT 29-MAR-1999 (first entry)  
 XX DE LS170 polypeptide sequence.  
 XX KW LS170 gene; in vivo imaging; lung disease; cancer; drug screening.  
 XX OS Homo sapiens.  
 XX SS Homo sapiens.  
 XX FH Key Location/Qualifiers  
 XX FT Peptide 17-DEC-1998.  
 XX FT Protein 11-JUN-1998; 98WO-US11601.  
 XX PF 11-JUN-1997; 97US-0049183.  
 XX PR 11-JUN-1997; 97US-0049183.  
 XX (ABBO ) ABBOTT LAB.  
 XX Billing-Medel LA, Cohen W, Colpitts TL, Friedman PJ:  
 PI Gordon J, Grandosien, Hodges SC, Klass WR, Kratochvil JB:  
 PI Roberts-Rapp L, Russell JC, Stroupe SD:  
 XX WPI: 1999-060335/05.  
 DR N-PSDB: AAX08093.



CC The invention relates to human nucleic acids (AA157798-AA161369) and  
 CC the encoded polypeptides (AA38642-AA42213) with neurotrophic,  
 CC immunosuppressant and cyostatic activity. The polynucleotides are useful  
 CC in gene therapy. A composition containing a polypeptide or polynucleotide  
 CC system, such as peripheral neuropathies, central nervous system diseases  
 CC such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 CC lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 CC utilisation of the activities such as: Immune system suppression,  
 CC Activin/inhibin activity, chemocytic/chemokinetic activity, haemostatic  
 CC and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 CC assays for receptor activity, arthritis and inflammation, leukaemias and  
 CC C.N.S disorders.  
 CC Note: The sequence data for this patent did not form part of the printed  
 CC specification.

XX Sequence 256 AA:

Query Match 100.0%; Score 33; DB 22; Length 256;  
 Best Local Similarity 100.0%; Pctd No. 3 Be+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGCTGC 6  
 |||||  
 Db 80 999tsg 85

# RESULT 12

AA897366 standard; protein: 256 AA.

XX AAB97366:

XX 15-ANG-2001 (first entry)

XX Human LUNX protein.

XX LUNX; human; cancer; micrometastatic cancer.

XX Homo sapiens.

XX JF200107872-A.

XX 27-MAR-2001.

XX 07-SEP-1999; 99JP-0253186.

XX 07-SEP-1999; 99JP-0253186.

XX (SMAK) OTSUKA PHARM CO LTD.

XX WPI: 2001-313367/33.

XX N-PSDB: AA827756, AA827757.

XX Polynucleotide encoding LUNX gene product useful for the detection of  
 XX cancer especially micrometastatic cancer.

XX Claim 1: Page 25; 30pp; Japanese.

XX This invention relates to the human LUNX protein and the polynucleotide  
 XX sequence encoding it. The invention includes a vector containing a LUNX  
 XX polynucleotide, a host cell transformed with the vector, and an antibody  
 XX that binds to LUNX. The gene can be used for cancer diagnosis and  
 XX diagnosis of micrometastatic cancer. The invention also includes a  
 XX gene product. The present sequence represents human LUNX.

XX Sequence 256 AA:

Query Match 100.0%; Score 33; DB 22; Length 256;  
 Best Local Similarity 100.0%; Pctd. No. 3 Be+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GGCTGC 6  
 |||||  
 Db 80 999tsg 85

# RESULT 13

AA441507 standard; Protein: 264 AA.

XX AA441507:

XX 22-OCT-2001 (first entry)

XX Human polypeptide SEQ ID NO 6438.

XX Human; neurotrophic; immunosuppressant; cyostatic; gene therapy; cancer;  
 XX peripheral nervous system, neuropathy; central nervous system;  
 XX Alzheimer's; lateral sclerosis; Shy-Drager Syndrome; chemocytic;  
 XX chemokinetic; thrombolytic; drug screening; arthritis; inflammation;  
 XX leukaemia.

XX Homo sapiens.

XX W020015312-A1.

XX 26-JUL-2001.

XX 26-DEC-2000; 2000MO-US34263.

XX 21-JAN-2000; 2000US-0488725.

XX 25-APR-2000; 2000US-052817.

XX 19-JUL-2000; 2000US-052817.

XX 03-AUG-2000; 2000US-0620312.

XX 14-SEP-2000; 2000US-0653450.

XX 19-OCT-2000; 2000US-0663191.

XX 29-NOV-2000; 2000US-0693036.

XX (HSE-) HSEQ INC.

XX Tang Y., Liu C., Asundi V., Chen R., Ma Y., Qian XB, Ren F., Wang D;  
 XX Wang J., Wang Z., Wehrman T., Xu C., Xue AJ, Yang Y., Zhang J;  
 XX Zhao OH, Zhou P, Goodrich R, Dermanac RT;

XX WPI: 2001-442253/47.

XX N-PSDB: AA160663.

XX Novel nucleic acids and polypeptides, useful for treating disorders  
 XX such as central nervous system injuries.

XX Example 2: SEQ ID NO 6438; 10078pp; English.

XX The invention relates to human nucleic acids (AA157798-AA161369) and  
 XX the encoded polypeptides (AA38642-AA42213) with neurotrophic,  
 XX immunosuppressant and cyostatic activity. The polynucleotides are useful  
 XX in gene therapy. A composition containing a polypeptide or polynucleotide  
 XX system, such as peripheral neuropathies, central nervous system diseases,  
 XX such as Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 XX utilisation of the activities such as: Immune system suppression,  
 XX Activin/inhibin activity, chemocytic/chemokinetic activity, haemostatic  
 XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,  
 XX assays for receptor activity, arthritis and inflammation, leukaemias and  
 XX C.N.S disorders.

XX Note: The sequence data for this patent did not form part of the printed  
 XX specification.

XX Sequence 264 AA:

Query Match 100.0%; Score 33; DB 22; Length 264;  
 Best Local Similarity 100.0%; Pctd No. 46402;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGCCTSG 6  
 |||||  
 Db 88 99CTSG 93

RESULT 14  
 AAM41789 standard; Protein: 272 AA.  
 ID AAM41789  
 AC AAM41789;  
 DM 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 6720.  
 DE Human polypeptide SEQ ID NO 6720.  
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM chondrocytic; lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chondrocytic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.  
 XX Homo sapiens.  
 OS  
 XX MO20015312-AL.  
 XX 26-JUL-2001.  
 XX 26-DEC-2000; 2000MO-US34263.  
 PF 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 19-JUL-2000; 2000US-0558042.  
 PR 03-AUG-2000; 2000US-0552451.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HSE-) HSESEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xu AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 DR NPI: 2001-442253/47.  
 DR N-PSDB: AA160945.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 PS Example 2; SEQ ID NO 6720; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and  
 the encoded polypeptides (AAM3662-AAM4213) with noctropic,  
 immunosuppressive, cytostatic activity. The polynucleotides are useful  
 in gene therapy. A composition containing the polynucleotides or polypeptide  
 system, such as peripheral nervous injuries, peripheral neuropathy and  
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 activation of the activities such as: immune system suppression,  
 chondrocytic, thrombolytic, drug screening, arthritis, inflammation,  
 leukaemia and chemotactic activity, cancer diagnosis, haemostatic  
 assays for receptor activity, arthritis and inflammation, leukaemia and  
 C.N.S disorders.

Note: The sequence data for this patent did not form part of the printed  
 CC specification.  
 CC  
 XX Sequence 272 AA;  
 S0

Query Match 100.0%; Score 33; DB 22; Length 272;  
 Best Local Similarity 100.0%; Pctd No. 46402;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Y 1 GGCCTSG 6  
 |||||  
 Db 123 99CTSG 128

RESULT 15  
 AAM41790 standard; Protein: 272 AA.  
 ID AAM41790  
 AC AAM41790;  
 DM 22-OCT-2001 (first entry)  
 XX Human polypeptide SEQ ID NO 6721.  
 DE Human polypeptide SEQ ID NO 6721.  
 XX Human; noctropic; immunosuppressant; cytostatic; gene therapy; cancer;  
 KM peripheral nervous system; neuropathy; central nervous system; CNS;  
 KM Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;  
 KM chondrocytic; lateral sclerosis; Shy-Drager Syndrome; chemotactic;  
 KM chondrocytic; thrombolytic; drug screening; arthritis; inflammation;  
 KM leukaemia.  
 XX Homo sapiens.  
 OS  
 XX MO20015312-AL.  
 XX 26-JUL-2001.  
 XX 26-DEC-2000; 2000MO-US34263.  
 PF 21-JAN-2000; 2000US-0488725.  
 PR 25-APR-2000; 2000US-0552317.  
 PR 19-JUL-2000; 2000US-0558042.  
 PR 03-AUG-2000; 2000US-0653450.  
 PR 14-SEP-2000; 2000US-0662191.  
 PR 19-OCT-2000; 2000US-0693036.  
 PR 29-NOV-2000; 2000US-0727344.  
 XX (HSE-) HSESEQ INC.  
 XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;  
 PI Wang J, Wang Z, Wehrman T, Xu C, Xu AJ, Yang Y, Zhang J;  
 PI Zhao QA, Zhou P, Goodrich R, Drmanac RT;  
 DR NPI: 2001-442253/47.  
 DR N-PSDB: AA160946.  
 XX Novel nucleic acids and polypeptides, useful for treating disorders  
 PT such as central nervous system injuries -  
 PS Example 2; SEQ ID NO 6721; 10078pp; English.

The invention relates to human nucleic acids (AA157798-AA161369) and  
 the encoded polypeptides (AAM3662-AAM4213) with noctropic,  
 immunosuppressive, cytostatic activity. The polynucleotides are useful  
 in gene therapy. A composition containing the polynucleotides or polypeptide  
 system, such as peripheral nervous injuries, peripheral neuropathy and  
 Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic  
 lateral sclerosis, and Shy-Drager Syndrome. Other uses include the  
 activation of the activities such as: immune system suppression,  
 chondrocytic, thrombolytic, drug screening, arthritis, inflammation,  
 leukaemia and chemotactic activity, cancer diagnosis, haemostatic  
 assays for receptor activity, arthritis and inflammation, leukaemia and  
 C.N.S disorders.



[illegible]

[illegible]

Query Match	Best Local	Similarity	100.0%	Score 33	DB 21	Length 300
Matches	6	Conservative	0	Mismatches	0	Gaps
Qy	1	GGCTSG	6			
Db	11	gggltsg	16			
RESULT 17						
ID	AA01576	AG11576 standard; Protein: 302 AA.				
XX	XX	AG11576:				
XX	XX	17-OCT-2000 (first entry)				
XX	XX	Arabidopsis thaliana protein fragment S50 ID NO: 10344.				
XX	XX	Protein identification, signal transduction pathway; metabolic pathway;				
XX	XX	hybridisation assay; genetic mapping; gene expression control; promoter;				
XX	XX	termination sequence.				
XX	XX	Arabidopsis thaliana.				
XX	XX	EP1033405-A2.				
XX	XX	06-SEP-2000.				
XX	XX	25-FEB-2000; 2000EP-0301439.				
XX	XX	25-FEB-1999; 9905-0121625.				
XX	XX	06-MAR-1999; 9905-0123180.				
XX	XX	09-MAR-1999; 9905-0125788.				
XX	XX	23-MAR-1999; 9905-0125788.				
XX	XX	25-MAR-1999; 9905-0126264.				
XX	XX	29-MAR-1999; 9905-0126785.				
XX	XX	01-APR-1999; 9905-0127462.				
XX	XX	06-APR-1999; 9905-0128234.				
XX	XX	11-APR-1999; 9905-0129645.				
XX	XX	19-APR-1999; 9905-0130444.				
XX	XX	21-APR-1999; 9905-0130444.				
XX	XX	23-APR-1999; 9905-0130510.				
XX	XX	23-APR-1999; 9905-0130891.				
XX	XX	28-APR-1999; 9905-0131444.				
XX	XX	30-APR-1999; 9905-0132048.				
XX	XX	04-MAY-1999; 9905-0132407.				
XX	XX	04-MAY-1999; 9905-0132468.				
XX	XX	05-MAY-1999; 9905-0132486.				
XX	XX	06-MAY-1999; 9905-0132487.				
XX	XX	06-MAY-1999; 9905-0132487.				
XX	XX	07-MAY-1999; 9905-0132863.				
XX	XX	11-MAY-1999; 9905-0134256.				
XX	XX	14-MAY-1999; 9905-0134218.				
XX	XX	14-MAY-1999; 9905-0134219.				
XX	XX	14-MAY-1999; 9905-0134221.				
XX	XX	14-MAY-1999; 9905-0134221.				
XX	XX	18-MAY-1999; 9905-0134768.				
XX	XX	19-MAY-1999; 9905-0134941.				
XX	XX	20-MAY-1999; 9905-0135124.				
XX	XX	21-MAY-1999; 9905-0135353.				
XX	XX	24-MAY-1999; 9905-0135629.				
XX	XX	25-MAY-1999; 9905-0136021.				
XX	XX	25-MAY-1999; 9905-0136982.				
XX	XX	28-MAY-1999; 9905-0137252.				
XX	XX	01-JUN-1999; 9905-0137252.				
XX	XX	03-JUN-1999; 9905-0137558.				
XX	XX	04-JUN-1999; 9905-0137502.				
XX	XX	07-JUN-1999; 9905-0137724.				

PR	08-JUN-1999:	99US-0138094:	PR	09-AUG-1999:	99US-0147933:	PR	09-AUG-1999:	99US-0147933:	PR	09-AUG-1999:	99US-0147933:
PR	10-JUN-1999:	99US-0138540:	PR	10-AUG-1999:	99US-0148171:	PR	10-AUG-1999:	99US-0148171:	PR	10-AUG-1999:	99US-0148171:
PR	14-JUN-1999:	99US-0139119:	PR	11-AUG-1999:	99US-0148319:	PR	11-AUG-1999:	99US-0148319:	PR	11-AUG-1999:	99US-0148319:
PR	16-JUN-1999:	99US-0139452:	PR	12-AUG-1999:	99US-0148441:	PR	12-AUG-1999:	99US-0148441:	PR	12-AUG-1999:	99US-0148441:
PR	16-JUN-1999:	99US-0139453:	PR	13-AUG-1999:	99US-0148583:	PR	13-AUG-1999:	99US-0148583:	PR	13-AUG-1999:	99US-0148583:
PR	17-JUN-1999:	99US-0139452:	PR	13-AUG-1999:	99US-0148688:	PR	13-AUG-1999:	99US-0148688:	PR	13-AUG-1999:	99US-0148688:
PR	18-JUN-1999:	99US-0139454:	PR	16-AUG-1999:	99US-0149368:	PR	16-AUG-1999:	99US-0149368:	PR	16-AUG-1999:	99US-0149368:
PR	18-JUN-1999:	99US-0139456:	PR	17-AUG-1999:	99US-0149759:	PR	17-AUG-1999:	99US-0149759:	PR	17-AUG-1999:	99US-0149759:
PR	18-JUN-1999:	99US-0139457:	PR	18-AUG-1999:	99US-0149726:	PR	18-AUG-1999:	99US-0149726:	PR	18-AUG-1999:	99US-0149726:
PR	18-JUN-1999:	99US-0139458:	PR	20-AUG-1999:	99US-0149723:	PR	20-AUG-1999:	99US-0149723:	PR	20-AUG-1999:	99US-0149723:
PR	18-JUN-1999:	99US-0139459:	PR	20-AUG-1999:	99US-0148593:	PR	20-AUG-1999:	99US-0148593:	PR	20-AUG-1999:	99US-0148593:
PR	18-JUN-1999:	99US-0139460:	PR	23-AUG-1999:	99US-0149930:	PR	23-AUG-1999:	99US-0149930:	PR	23-AUG-1999:	99US-0149930:
PR	18-JUN-1999:	99US-0139463:	PR	25-AUG-1999:	99US-0150566:	PR	25-AUG-1999:	99US-0150566:	PR	25-AUG-1999:	99US-0150566:
PR	18-JUN-1999:	99US-0139750:	PR	26-AUG-1999:	99US-0150884:	PR	26-AUG-1999:	99US-0150884:	PR	26-AUG-1999:	99US-0150884:
PR	18-JUN-1999:	99US-0139763:	PR	27-AUG-1999:	99US-0151065:	PR	27-AUG-1999:	99US-0151065:	PR	27-AUG-1999:	99US-0151065:
PR	21-JUN-1999:	99US-0139817:	PR	27-AUG-1999:	99US-0151066:	PR	27-AUG-1999:	99US-0151066:	PR	27-AUG-1999:	99US-0151066:
PR	22-JUN-1999:	99US-0139899:	PR	30-AUG-1999:	99US-0151080:	PR	30-AUG-1999:	99US-0151080:	PR	30-AUG-1999:	99US-0151080:
PR	23-JUN-1999:	99US-0140353:	PR	30-AUG-1999:	99US-0151478:	PR	30-AUG-1999:	99US-0151478:	PR	30-AUG-1999:	99US-0151478:
PR	23-JUN-1999:	99US-0140625:	PR	01-SEP-1999:	99US-0151930:	PR	01-SEP-1999:	99US-0151930:	PR	01-SEP-1999:	99US-0151930:
PR	28-JUN-1999:	99US-0140823:	PR	07-SEP-1999:	99US-0153363:	PR	07-SEP-1999:	99US-0153363:	PR	07-SEP-1999:	99US-0153363:
PR	29-JUN-1999:	99US-0140951:	PR	10-SEP-1999:	99US-0153070:	PR	10-SEP-1999:	99US-0153070:	PR	10-SEP-1999:	99US-0153070:
PR	30-JUN-1999:	99US-0141287:	PR	13-SEP-1999:	99US-0153758:	PR	13-SEP-1999:	99US-0153758:	PR	13-SEP-1999:	99US-0153758:
PR	01-JUL-1999:	99US-0141842:	PR	15-SEP-1999:	99US-0154018:	PR	15-SEP-1999:	99US-0154018:	PR	15-SEP-1999:	99US-0154018:
PR	02-JUL-1999:	99US-0142154:	PR	16-SEP-1999:	99US-0154973:	PR	16-SEP-1999:	99US-0154973:	PR	16-SEP-1999:	99US-0154973:
PR	02-JUL-1999:	99US-0142905:	PR	20-SEP-1999:	99US-0155293:	PR	20-SEP-1999:	99US-0155293:	PR	20-SEP-1999:	99US-0155293:
PR	08-JUL-1999:	99US-0142803:	PR	23-SEP-1999:	99US-0155119:	PR	23-SEP-1999:	99US-0155119:	PR	23-SEP-1999:	99US-0155119:
PR	09-JUL-1999:	99US-0142920:	PR	24-SEP-1999:	99US-0155486:	PR	24-SEP-1999:	99US-0155486:	PR	24-SEP-1999:	99US-0155486:
PR	12-JUL-1999:	99US-0142977:	PR	28-SEP-1999:	99US-0156458:	PR	28-SEP-1999:	99US-0156458:	PR	28-SEP-1999:	99US-0156458:
PR	13-JUL-1999:	99US-0143542:	PR	29-SEP-1999:	99US-0156596:	PR	29-SEP-1999:	99US-0156596:	PR	29-SEP-1999:	99US-0156596:
PR	14-JUL-1999:	99US-0143624:	PR	04-OCT-1999:	99US-0157117:	PR	04-OCT-1999:	99US-0157117:	PR	04-OCT-1999:	99US-0157117:
PR	15-JUL-1999:	99US-0144002:	PR	05-OCT-1999:	99US-0157865:	PR	05-OCT-1999:	99US-0157865:	PR	05-OCT-1999:	99US-0157865:
PR	16-JUL-1999:	99US-0144086:	PR	07-OCT-1999:	99US-0158029:	PR	07-OCT-1999:	99US-0158029:	PR	07-OCT-1999:	99US-0158029:
PR	19-JUL-1999:	99US-0144325:	PR	08-OCT-1999:	99US-0158232:	PR	08-OCT-1999:	99US-0158232:	PR	08-OCT-1999:	99US-0158232:
PR	19-JUL-1999:	99US-0144331:	PR	12-OCT-1999:	99US-0158369:	PR	12-OCT-1999:	99US-0158369:	PR	12-OCT-1999:	99US-0158369:
PR	19-JUL-1999:	99US-0144332:	PR	13-OCT-1999:	99US-0159293:	PR	13-OCT-1999:	99US-0159293:	PR	13-OCT-1999:	99US-0159293:
PR	19-JUL-1999:	99US-0144333:	PR	13-OCT-1999:	99US-0159294:	PR	13-OCT-1999:	99US-0159294:	PR	13-OCT-1999:	99US-0159294:
PR	19-JUL-1999:	99US-0144334:	PR	13-OCT-1999:	99US-0159329:	PR	13-OCT-1999:	99US-0159329:	PR	13-OCT-1999:	99US-0159329:
PR	19-JUL-1999:	99US-0144335:	PR	14-OCT-1999:	99US-0159330:	PR	14-OCT-1999:	99US-0159330:	PR	14-OCT-1999:	99US-0159330:
PR	20-JUL-1999:	99US-0144632:	PR	14-OCT-1999:	99US-0159331:	PR	14-OCT-1999:	99US-0159331:	PR	14-OCT-1999:	99US-0159331:
PR	20-JUL-1999:	99US-0144884:	PR	14-OCT-1999:	99US-0159637:	PR	14-OCT-1999:	99US-0159637:	PR	14-OCT-1999:	99US-0159637:
PR	21-JUL-1999:	99US-0144814:	PR	14-OCT-1999:	99US-0159638:	PR	14-OCT-1999:	99US-0159638:	PR	14-OCT-1999:	99US-0159638:
PR	21-JUL-1999:	99US-0145086:	PR	18-OCT-1999:	99US-0159584:	PR	18-OCT-1999:	99US-0159584:	PR	18-OCT-1999:	99US-0159584:
PR	21-JUL-1999:	99US-0145088:	PR	21-OCT-1999:	99US-0160741:	PR	21-OCT-1999:	99US-0160741:	PR	21-OCT-1999:	99US-0160741:
PR	22-JUL-1999:	99US-0145089:	PR	21-OCT-1999:	99US-0160768:	PR	21-OCT-1999:	99US-0160768:	PR	21-OCT-1999:	99US-0160768:
PR	22-JUL-1999:	99US-0145089:	PR	21-OCT-1999:	99US-0160770:	PR	21-OCT-1999:	99US-0160770:	PR	21-OCT-1999:	99US-0160770:
PR	22-JUL-1999:	99US-0145192:	PR	21-OCT-1999:	99US-0160814:	PR	21-OCT-1999:	99US-0160814:	PR	21-OCT-1999:	99US-0160814:
PR	23-JUL-1999:	99US-0145145:	PR	21-OCT-1999:	99US-0160815:	PR	21-OCT-1999:	99US-0160815:	PR	21-OCT-1999:	99US-0160815:
PR	23-JUL-1999:	99US-0145218:	PR	22-OCT-1999:	99US-0160981:	PR	22-OCT-1999:	99US-0160981:	PR	22-OCT-1999:	99US-0160981:
PR	23-JUL-1999:	99US-0145222:	PR	22-OCT-1999:	99US-0160989:	PR	22-OCT-1999:	99US-0160989:	PR	22-OCT-1999:	99US-0160989:
PR	26-JUL-1999:	99US-0145276:	PR	22-OCT-1999:	99US-0160989:	PR	22-OCT-1999:	99US-0160989:	PR	22-OCT-1999:	99US-0160989:
PR	27-JUL-1999:	99US-0145913:	PR	25-OCT-1999:	99US-0161405:	PR	25-OCT-1999:	99US-0161405:	PR	25-OCT-1999:	99US-0161405:
PR	27-JUL-1999:	99US-0145913:	PR	25-OCT-1999:	99US-0161406:	PR	25-OCT-1999:	99US-0161406:	PR	25-OCT-1999:	99US-0161406:
PR	28-JUL-1999:	99US-0145951:	PR	26-OCT-1999:	99US-0161360:	PR	26-OCT-1999:	99US-0161360:	PR	26-OCT-1999:	99US-0161360:
PR	28-AUG-1999:	99US-0146388:	PR	26-OCT-1999:	99US-0161361:	PR	26-OCT-1999:	99US-0161361:	PR	26-OCT-1999:	99US-0161361:
PR	02-AUG-1999:	99US-0146388:	PR	28-OCT-1999:	99US-0161360:	PR	28-OCT-1999:	99US-0161360:	PR	28-OCT-1999:	99US-0161360:
PR	03-AUG-1999:	99US-0147038:	PR	28-OCT-1999:	99US-0161360:	PR	28-OCT-1999:	99US-0161360:	PR	28-OCT-1999:	99US-0161360:
PR	04-AUG-1999:	99US-0147302:	PR	28-OCT-1999:	99US-0161993:	PR	28-OCT-1999:	99US-0161993:	PR	28-OCT-1999:	99US-0161993:
PR	05-AUG-1999:	99US-0147792:	PR	29-OCT-1999:	99US-0162142:	PR	29-OCT-1999:	99US-0162142:	PR	29-OCT-1999:	99US-0162142:
PR	05-AUG-1999:	99US-0147760:	PR			PR			PR		
PR	06-AUG-1999:	99US-0147703:	PR			PR			PR		
PR	06-AUG-1999:	99US-0147416:	PR			PR			PR		

Query Match 100.0% Score 33: DB 21: Length 302:  
 Best Local Similarity 100.0% Pred. No. 4.4e+02:  
 Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

0Y 1 GCGTSG 6  
Db 13 999459 18  
RESULT 18  
AA01575  
ID AG01575 standard; Protein: 303 AA.  
AC AG01575:  
XX  
DT 17-OCT-2000 (first entry)  
XX  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10343.  
XX  
KW Protein identification; signal transduction pathway; metabolic pathway;  
XX Hybridization; mass genetic mapping; gene expression control; promoter;  
XX Termination sequence.  
OS Arabidopsis thaliana.  
XX  
XX  
XX EPI033405-A2.  
PD 06-SEP-2000.  
XX  
PF 25-FEB-2000; 2000EP-0301439.  
XX  
XX 25-FEB-1999; 9905-0121825.  
XX 03-MAR-1999; 9905-0123180.  
XX 23-MAR-1999; 9905-0123588.  
XX 25-MAR-1999; 9905-0126264.  
XX 29-MAR-1999; 9905-0126785.  
XX 01-APR-1999; 9905-0127462.  
XX 06-APR-1999; 9905-0128234.  
XX 08-APR-1999; 9905-0128814.  
XX 15-APR-1999; 9905-0129845.  
XX 21-APR-1999; 9905-0130440.  
XX 23-APR-1999; 9905-0130510.  
XX 23-APR-1999; 9905-0130891.  
XX 28-APR-1999; 9905-0131449.  
XX 30-APR-1999; 9905-0132048.  
XX 04-MAY-1999; 9905-0132407.  
XX 05-MAY-1999; 9905-0132484.  
XX 06-MAY-1999; 9905-0132486.  
XX 06-MAY-1999; 9905-0132487.  
XX 07-MAY-1999; 9905-0132653.  
XX 11-MAY-1999; 9905-0134456.  
XX 13-MAY-1999; 9905-0134418.  
XX 14-MAY-1999; 9905-0134429.  
XX 14-MAY-1999; 9905-0134470.  
XX 18-MAY-1999; 9905-0134768.  
XX 19-MAY-1999; 9905-0134941.  
XX 20-MAY-1999; 9905-0135124.  
XX 22-MAY-1999; 9905-0135253.  
XX 24-MAY-1999; 9905-0136021.  
XX 25-MAY-1999; 9905-0136023.  
XX 27-MAY-1999; 9905-0136593.  
XX 28-MAY-1999; 9905-0136782.  
XX 01-JUN-1999; 9905-0137222.  
XX 03-JUN-1999; 9905-0137528.  
XX 04-JUN-1999; 9905-0137502.  
XX 05-JUN-1999; 9905-0137724.  
XX 08-JUN-1999; 9905-0138540.  
XX 10-JUN-1999; 9905-0138540.  
XX 10-JUN-1999; 9905-0138847.  
XX 14-JUN-1999; 9905-0139119.  
XX 16-JUN-1999; 9905-0139453.  
XX 17-JUN-1999; 9905-0139453.  
XX 9905-0139492.

PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139464.  
PR 18-JUN-1999; 9905-0139465.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139763.  
PR 21-JUN-1999; 9905-0139817.  
PR 23-JUN-1999; 9905-0139859.  
PR 23-JUN-1999; 9905-0140054.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 01-JUL-1999; 9905-0141942.  
PR 01-JUL-1999; 9905-0142054.  
PR 02-JUL-1999; 9905-0142390.  
PR 06-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 11-JUL-1999; 9905-0142977.  
PR 12-JUL-1999; 9905-0143242.  
PR 14-JUL-1999; 9905-0143542.  
PR 15-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144333.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144684.  
PR 21-JUL-1999; 9905-0144816.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145345.  
PR 23-JUL-1999; 9905-0145718.  
PR 26-JUL-1999; 9905-0145776.  
PR 27-JUL-1999; 9905-0145913.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 27-JUL-1999; 9905-0145931.  
PR 02-AUG-1999; 9905-0145932.  
PR 02-AUG-1999; 9905-0145988.  
PR 02-AUG-1999; 9905-0146388.  
PR 03-AUG-1999; 9905-0146389.  
PR 04-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 05-AUG-1999; 9905-0147209.  
PR 05-AUG-1999; 9905-0147199.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147416.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147955.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 12-AUG-1999; 9905-0148341.  
PR 13-AUG-1999; 9905-0148355.  
PR 13-AUG-1999; 9905-0148684.

PR	16-AUG-1999	9905-0149368
PR	17-AUG-1999	9905-0149175
PR	18-AUG-1999	9905-0149426
PR	19-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149723
PR	20-AUG-1999	9905-0149902
PR	23-AUG-1999	9905-0149930
PR	23-AUG-1999	9905-0150566
PR	25-AUG-1999	9905-0150864
PR	26-AUG-1999	9905-0151065
PR	27-AUG-1999	9905-0151066
PR	27-AUG-1999	9905-0151103
PR	27-AUG-1999	9905-0151103
PR	31-AUG-1999	9905-0151438
PR	01-SEP-1999	9905-0151930
PR	07-SEP-1999	9905-0152653
PR	10-SEP-1999	9905-0153070
PR	13-SEP-1999	9905-0153758
PR	15-SEP-1999	9905-0154018
PR	16-SEP-1999	9905-0154059
PR	16-SEP-1999	9905-0154139
PR	20-SEP-1999	9905-0155139
PR	23-SEP-1999	9905-0155586
PR	23-SEP-1999	9905-0155559
PR	28-SEP-1999	9905-0156458
PR	29-SEP-1999	9905-0156596
PR	04-OCT-1999	9905-0157117
PR	05-OCT-1999	9905-0157753
PR	05-OCT-1999	9905-0157878
PR	06-OCT-1999	9905-0158025
PR	06-OCT-1999	9905-0158232
PR	12-OCT-1999	9905-0158369
PR	13-OCT-1999	9905-0159293
PR	13-OCT-1999	9905-0159294
PR	13-OCT-1999	9905-0159295
PR	14-OCT-1999	9905-0159329
PR	14-OCT-1999	9905-0159330
PR	14-OCT-1999	9905-0159331
PR	14-OCT-1999	9905-0159637
PR	14-OCT-1999	9905-0159658
PR	18-OCT-1999	9905-0159584
PR	21-OCT-1999	9905-0160741
PR	21-OCT-1999	9905-0160767
PR	21-OCT-1999	9905-0160768
PR	21-OCT-1999	9905-0160770
PR	21-OCT-1999	9905-0160814
PR	21-OCT-1999	9905-0160814
PR	22-OCT-1999	9905-0160980
PR	22-OCT-1999	9905-0160981
PR	25-OCT-1999	9905-0160989
PR	25-OCT-1999	9905-0161404
PR	25-OCT-1999	9905-0161405
PR	25-OCT-1999	9905-0161406
PR	26-OCT-1999	9905-0161380
PR	26-OCT-1999	9905-0161381
PR	26-OCT-1999	9905-0161381
PR	28-OCT-1999	9905-0161920
PR	28-OCT-1999	9905-0161992
PR	29-OCT-1999	9905-0161993
PR	29-OCT-1999	9905-0162142

```

Query Match Similarity 100.0% Score 33: DB 21: Length 303:
Best Local Similarity 100.0% Pried. No. 4.4e+02:
Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:
QY 1 GGGTSG 6
    111111
Db 14 gggctsg 19

```

ID	AAV32858 standard; Protein: 348 AA.
AC	AAV32858:
DR	01-NOV-1999 (first entry)
XX	
DE	Human heat shock protein homologue 1 (HSPH-1) protein sequence.
XX	
KW	Human heat shock protein homologue; HSPH-1; HSPH-2; diagnosis; cancer;
XN	Inflammation; AIDS; Crohn's disease; allergy; therapy; drug screening.
OS	Homo sapiens.
PX	US9545287-A.
PD	31-AUG-1999.
PE	19-NOV-1997; 9705-0974546.
PR	19-NOV-1997; 9705-0974546.
PS	(INCYTE - INCYTE PHARM INC.
XX	
PI	Corley MC, Hillman JL, Lal P, Shah P;
DR	WP1: 1999-526253/44.
DR	N-PSDB: AAT11066.
PT	New molecular techniques useful for diagnosis, prevention and
XX	treatment of cancer and inflammation
PS	Claim 1: Fig 1: 35pp: English.
XX	
CC	This sequence is the human heat shock protein homologue, HSPH-1 of
CC	the invention. The HSPH-1 polynucleotide was obtained from the BRAT1721
CC	cDNA library isolated from the BRS2M018 cDNA library constructed from
CC	dissolved breast tissue. The HSPH polynucleotides and their complements
CC	are useful for diagnosis of conditions or disorders associated with HSPH
CC	expression. HSPH complements are also useful as antagonists for
CC	prevention or treatment of cancer e.g. adenocarcinoma, breast and skin,
CC	and inflammation associated with e.g. AIDS, Crohn's disease and aging
CC	allergies. HSPH-1 and HSPH-2 are useful as antisense agents against
CC	libraries of compounds used in the study of HSPH-2 nucleotide sequences, and
CC	e.g. promoters and regulatory elements. Vectors containing the DNA
CC	probes.
XX	
SO	Sequence 348 AA:
XX	
QY	Query Mntch 100.0%; Score 33; DB 20; Length 348:
DB	Heat Local Similarity 100.0%; Pred.No. 5e+02;
XX	Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
XX	Cy 1 GCGTSG 6
XX	
XX	DB 75 GGAGSg 80
XX	
RESULT 20	
ID	AAB94509 standard; Protein: 348 AA.
AC	AAB94509:
DR	26-JUN-2003 (first entry)
XX	
DE	Human protein sequence SEQ ID NO:15218.
XX	
XX	Human; primer; detection; diagnosis; antisense therapy; gene therapy.



[illegible]

PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155487.  
PR 28-SEP-1999: 99US-0155488.  
PR 29-SEP-1999: 99US-0155489.  
PR 04-OCT-1999: 99US-0155536.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157865.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158156.  
PR 12-OCT-1999: 99US-0158293.  
PR 13-OCT-1999: 99US-0159294.  
PR 13-OCT-1999: 99US-0159295.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159384.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160741.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160768.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160815.  
PR 22-OCT-1999: 99US-0160980.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0160989.  
PR 22-OCT-1999: 99US-0161404.  
PR 22-OCT-1999: 99US-0161405.  
PR 22-OCT-1999: 99US-0161406.  
PR 26-OCT-1999: 99US-0161350.  
PR 26-OCT-1999: 99US-0161361.  
PR 26-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161993.  
PR 28-OCT-1999: 99US-0161993.  
PR 28-OCT-1999: 99US-0162142.  
PR 29-OCT-1999: 99US-0162142.

## Query Match

100.0%; Score 33; DB 21; Length 499.

Best Local Similarity: 100.0%; Prod NO: 6; gaps: 0;

Matches: 6; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0;

Oy 1 GCGTGC 6  
Db 25 999cag 30

## RESULT 23

AAC48191  
ID AAC48191 standard: Protein: 499 AA.

AAC48191:

18-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment seq ID NO: 60832.

DE Protein Identification: signal transduction pathway: metabolic pathway;  
KW Protein Identification assay: genetic mapping: gene expression control: promoter;  
KW Termination sequence.

Arabidopsis thaliana.

EP1031405-A2.

06-SEP-2000.

25-FEB-2000: 2000EP-030139.

25-FEB-1999: 99US-0121825.  
PR 05-MAR-1999: 99US-0123180.  
PR 09-MAR-1999: 99US-0123548.

PR 23-MAR-1999: 99US-0125788.  
PR 25-MAR-1999: 99US-0126264.  
PR 29-MAR-1999: 99US-0126785.  
PR 01-APR-1999: 99US-0127462.  
PR 06-APR-1999: 99US-0128234.  
PR 08-APR-1999: 99US-0128714.  
PR 16-APR-1999: 99US-0130077.  
PR 19-APR-1999: 99US-0130449.  
PR 21-APR-1999: 99US-0130510.  
PR 23-APR-1999: 99US-0130891.  
PR 23-APR-1999: 99US-0131449.  
PR 26-APR-1999: 99US-0132048.  
PR 30-APR-1999: 99US-0132484.  
PR 04-MAY-1999: 99US-0132485.  
PR 06-MAY-1999: 99US-0132486.  
PR 06-MAY-1999: 99US-0132487.  
PR 07-MAY-1999: 99US-0132863.  
PR 14-MAY-1999: 99US-0134266.  
PR 14-MAY-1999: 99US-0134216.  
PR 14-MAY-1999: 99US-0134221.  
PR 14-MAY-1999: 99US-0134370.  
PR 18-MAY-1999: 99US-0134368.  
PR 18-MAY-1999: 99US-0134941.  
PR 20-MAY-1999: 99US-0135124.  
PR 21-MAY-1999: 99US-0135253.  
PR 24-MAY-1999: 99US-0136029.  
PR 25-MAY-1999: 99US-0136029.  
PR 27-MAY-1999: 99US-0136392.  
PR 28-MAY-1999: 99US-0136782.  
PR 01-JUN-1999: 99US-0137222.  
PR 03-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 04-JUN-1999: 99US-0138024.  
PR 08-JUN-1999: 99US-0138040.  
PR 10-JUN-1999: 99US-0138040.  
PR 10-JUN-1999: 99US-0138047.  
PR 14-JUN-1999: 99US-0139159.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 16-JUN-1999: 99US-0139454.  
PR 16-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139462.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139763.  
PR 21-JUN-1999: 99US-0139817.  
PR 23-JUN-1999: 99US-0139859.  
PR 23-JUN-1999: 99US-0140154.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 02-JUL-1999: 99US-0142525.  
PR 06-JUL-1999: 99US-0142390.  
PR 08-JUL-1999: 99US-0142803.  
PR 09-JUL-1999: 99US-0142920.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 13-JUL-1999: 99US-0143624.  
PR 15-JUL-1999: 99US-0144005.



PR	16	JULI	1993	9505	0144008	
PR	15	JULI	1993	9505	0144035	
PR	15	JULI	1993	9505	0144331	
PR	15	JULI	1993	9505	0144331	
PR	15	JULI	1993	9505	0144331	
PR	15	JULI	1993	9505	0144334	
PR	15	JULI	1993	9505	0144334	
PR	15	JULI	1993	9505	0144334	
PR	15	JULI	1993	9505	0144335	
PR	20	JULI	1993	9505	0144332	
PR	20	JULI	1993	9505	0144632	
PR	20	JULI	1993	9505	0144814	
PR	21	JULI	1993	9505	0144814	
PR	21	JULI	1993	9505	0144814	
PR	21	JULI	1993	9505	0144814	
PR	22	JULI	1993	9505	0145087	
PR	22	JULI	1993	9505	0145087	
PR	22	JULI	1993	9505	0145108	
PR	22	JULI	1993	9505	0145119	
PR	22	JULI	1993	9505	0145115	
PR	23	JULI	1993	9505	0145228	
PR	23	JULI	1993	9505	0145228	
PR	23	JULI	1993	9505	0145228	
PR	23	JULI	1993	9505	0145228	
PR	27	JULI	1993	9505	0145913	
PR	27	JULI	1993	9505	0145918	
PR	27	JULI	1993	9505	0145918	
PR	28	JULI	1993	9505	0145931	
PR	02	AUGI	1993	9505	0146386	
PR	02	AUGI	1993	9505	0146386	
PR	02	AUGI	1993	9505	0146389	
PR	03	AUGI	1993	9505	0147038	
PR	04	AUGI	1993	9505	0147302	
PR	04	AUGI	1993	9505	0147302	
PR	05	AUGI	1993	9505	0147192	
PR	05	AUGI	1993	9505	0147280	
PR	06	AUGI	1993	9505	0147303	
PR	06	AUGI	1993	9505	0147303	
PR	09	AUGI	1993	9505	0147415	
PR	09	AUGI	1993	9505	0147415	
PR	10	AUGI	1993	9505	0147935	
PR	10	AUGI	1993	9505	0148171	
PR	11	AUGI	1993	9505	0148311	
PR	12	AUGI	1993	9505	0148311	
PR	13	AUGI	1993	9505	0148684	
PR	13	AUGI	1993	9505	0148684	
PR	13	AUGI	1993	9505	0148684	
PR	13	AUGI	1993	9505	0148684	
PR	16	AUGI	1993	9505	0148426	
PR	20	AUGI	1993	9505	0149722	
PR	20	AUGI	1993	9505	0149723	
PR	20	AUGI	1993	9505	0149592	
PR	23	AUGI	1993	9505	0149592	
PR	23	AUGI	1993	9505	0149592	
PR	23	AUGI	1993	9505	0149592	
PR	26	AUGI	1993	9505	0150884	
PR	27	AUGI	1993	9505	0151065	
PR	27	AUGI	1993	9505	0151065	
PR	30	AUGI	1993	9505	0151080	
PR	30	AUGI	1993	9505	0151303	
PR	31	AUGI	1993	9505	0151438	
PR	01	SEPI	1993	9505	0151530	
PR	01	SEPI	1993	9505	0151530	
PR	10	SEPI	1993	9505	0151070	
PR	10	SEPI	1993	9505	0151070	
PR	11	SEPI	1993	9505	0151578	
PR	16	SEPI	1993	9505	0154039	
PR	16	SEPI	1993	9505	0154039	
PR	22	SEPI	1993	9505	0155139	
PR	22	SEPI				

PR	07-Oct-1999.	9905-0158029.
PR	08-Oct-1999.	9905-0158232.
PR	12-Oct-1999.	9905-0158359.
PR	13-Oct-1999.	9905-0159253.
PR	13-Oct-1999.	9905-0159284.
PR	13-Oct-1999.	9905-0159329.
PR	14-Oct-1999.	9905-0159330.
PR	14-Oct-1999.	9905-0159331.
PR	14-Oct-1999.	9905-0159637.
PR	14-Oct-1999.	9905-0159638.
PR	18-Oct-1999.	9905-0159584.
PR	21-Oct-1999.	9905-0160741.
PR	21-Oct-1999.	9905-0160761.
PR	21-Oct-1999.	9905-0160766.
PR	21-Oct-1999.	9905-0160814.
PR	21-Oct-1999.	9905-0160815.
PR	22-Oct-1999.	9905-0160980.
PR	22-Oct-1999.	9905-0160981.
PR	22-Oct-1999.	9905-0160989.
PR	25-Oct-1999.	9905-0161404.
PR	25-Oct-1999.	9905-0161405.
PR	25-Oct-1999.	9905-0161406.
PR	25-Oct-1999.	9905-0161360.
PR	26-Oct-1999.	9905-0161361.
PR	26-Oct-1999.	9905-0161920.
PR	28-Oct-1999.	9905-0161932.
PR	28-Oct-1999.	9905-0161933.
PR	29-Oct-1999.	9905-0162142.

```

Oy      1 CGCTGCG 6
Db      25 gqstg 30
          |||||
RESULT 24.
ID      AAB93590
AA      AAB93590 standard; Protein: 520 AA.
XX
XX      AAB93590.
XX
Df      26-JUN-2001 (first entry)
Dd      Human protein sequence SEQ ID NO:13016.
Dd      Human; primer: detection; diagnosis; antisense therapy; gene therapy.
XX
XX      Homo sapiens.
OS
XX      EP1074617-A2.
PN
PD      07-FEB-2001.
XX
XX      28-JUL-2000; 2000EP-0116126.
XX
XX      29-JUL-1999; 99JP-0248036.
PR      27-AUG-1999; 99JP-0300253.
PR      11-JAN-2000; 2000JP-0118776.
PR      02-MAY-2000; 2000JP-0183767.
PR      09-JUN-2000; 2000JP-0241899.
XX
XX      (HELI-) HELIX RES INST.
XX
XX      Ota T, Isegaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;
XX      Ishii S, Sugiyama T, Wakamatsu T, Negai K, Otsuki R.
XX      WPI: 2001-318749/34.
XX

```



PR 19-JUL-1999; 9905-0144333  
 PR 19-JUL-1999; 9905-0144334  
 PR 20-JUL-1999; 9905-0144335  
 PR 20-JUL-1999; 9905-0144336  
 PR 20-JUL-1999; 9905-0144337  
 PR 20-JUL-1999; 9905-0144338  
 PR 21-JUL-1999; 9905-0144339  
 PR 21-JUL-1999; 9905-0145086  
 PR 21-JUL-1999; 9905-0145087  
 PR 21-JUL-1999; 9905-0145088  
 PR 21-JUL-1999; 9905-0145089  
 PR 22-JUL-1999; 9905-0145090  
 PR 22-JUL-1999; 9905-0145192  
 PR 22-JUL-1999; 9905-0145193  
 PR 23-JUL-1999; 9905-0145145  
 PR 23-JUL-1999; 9905-0145218  
 PR 23-JUL-1999; 9905-0145224  
 PR 23-JUL-1999; 9905-0145276  
 PR 23-JUL-1999; 9905-0145277  
 PR 27-JUL-1999; 9905-0145518  
 PR 27-JUL-1999; 9905-0145519  
 PR 28-JUL-1999; 9905-0145551  
 PR 02-AUG-1999; 9905-0146386  
 PR 02-AUG-1999; 9905-0146387  
 PR 02-AUG-1999; 9905-0146388  
 PR 02-AUG-1999; 9905-0146389  
 PR 04-AUG-1999; 9905-0147198  
 PR 04-AUG-1999; 9905-0147202  
 PR 05-AUG-1999; 9905-0147192  
 PR 05-AUG-1999; 9905-0147260  
 PR 06-AUG-1999; 9905-0147303  
 PR 09-AUG-1999; 9905-0147416  
 PR 09-AUG-1999; 9905-0147493  
 PR 10-AUG-1999; 9905-0148135  
 PR 11-AUG-1999; 9905-0148310  
 PR 12-AUG-1999; 9905-0148341  
 PR 13-AUG-1999; 9905-0148365  
 PR 13-AUG-1999; 9905-0148684  
 PR 16-AUG-1999; 9905-0149368  
 PR 18-AUG-1999; 9905-0149172  
 PR 20-AUG-1999; 9905-0149723  
 PR 20-AUG-1999; 9905-0149729  
 PR 23-AUG-1999; 9905-0149802  
 PR 23-AUG-1999; 9905-0149930  
 PR 23-AUG-1999; 9905-0150566  
 PR 27-AUG-1999; 9905-0150884  
 PR 27-AUG-1999; 9905-0151066  
 PR 27-AUG-1999; 9905-0151080  
 PR 30-AUG-1999; 9905-0151303  
 PR 31-AUG-1999; 9905-0151438  
 PR 01-SEP-1999; 9905-0151930  
 PR 01-SEP-1999; 9905-0152363  
 PR 01-SEP-1999; 9905-0152364  
 PR 13-SEP-1999; 9905-0152758  
 PR 15-SEP-1999; 9905-0154018  
 PR 16-SEP-1999; 9905-0154039  
 PR 20-SEP-1999; 9905-0154779  
 PR 22-SEP-1999; 9905-0155139  
 PR 23-SEP-1999; 9905-0155486  
 PR 26-SEP-1999; 9905-0155659  
 PR 26-SEP-1999; 9905-0155659  
 PR 04-OCT-1999; 9905-0157517  
 PR 05-OCT-1999; 9905-0157531  
 PR 07-OCT-1999; 9905-0158029  
 PR 12-OCT-1999; 9905-0158432  
 PR 13-OCT-1999; 9905-0159293  
 PR 13-OCT-1999; 9905-0159294

PR 13-OCT-1999; 9905-0159295  
 PR 14-OCT-1999; 9905-0159329  
 PR 14-OCT-1999; 9905-0159330  
 PR 14-OCT-1999; 9905-0159331  
 PR 14-OCT-1999; 9905-0159633  
 PR 18-OCT-1999; 9905-0159584  
 PR 21-OCT-1999; 9905-0160741  
 PR 21-OCT-1999; 9905-0160767  
 PR 21-OCT-1999; 9905-0160768  
 PR 21-OCT-1999; 9905-0160770  
 PR 21-OCT-1999; 9905-0160814  
 PR 21-OCT-1999; 9905-0160833  
 PR 22-OCT-1999; 9905-0160843  
 PR 22-OCT-1999; 9905-0160981  
 PR 22-OCT-1999; 9905-0160988  
 PR 25-OCT-1999; 9905-0161404  
 PR 25-OCT-1999; 9905-0161406  
 PR 25-OCT-1999; 9905-0161359  
 PR 26-OCT-1999; 9905-0161361  
 PR 26-OCT-1999; 9905-0161361  
 PR 28-OCT-1999; 9905-0161392  
 PR 28-OCT-1999; 9905-0161993  
 PR 28-OCT-1999; 9905-0161993  
 PR 29-OCT-1999; 9905-0162142

Query Match 100.0% Score 33; DB 21; Length 552;  
 Best Local Similarity 100.0%  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTGC 6  
 DB 78 999tsg 83

RESULT 26  
 AAC48190  
 ID AAC48190 standard; Protein: 552 AA.  
 AC XX  
 AAC48190:  
 XX  
 18-OCT-2000 (first entry)  
 DE XX  
 Arabidopsis thaliana protein fragment SEQ ID NO: 60831.  
 XX  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 CO Arabidopsis thaliana.  
 XX  
 PV EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX  
 XX 25-FEB-2000; 2000EP-0301439.  
 XX  
 XX 25-FEB-1999; 9905-0121895.  
 XX 05-MAR-1999; 9905-0123160.  
 PR 09-MAR-1999; 9905-0123548.  
 PR 23-MAR-1999; 9905-0125788.  
 PR 25-MAR-1999; 9905-0126264.  
 PR 29-MAR-1999; 9905-0126795.  
 PR 01-APR-1999; 9905-0127462.  
 PR 08-APR-1999; 9905-0128274.  
 PR 08-APR-1999; 9905-0128845.  
 PR 16-APR-1999; 9905-0130077.  
 PR 19-APR-1999; 9905-0130077.  
 PR 21-APR-1999; 9905-0130449.  
 PR 23-APR-1999; 9905-0130510.  
 PR 28-APR-1999; 9905-0130891.  
 PR 28-APR-1999; 9905-0131449.

PR 30-APR-1999: 9905-01324048-  
PR 30-APR-1999: 9905-0132407-  
PR 04-MAY-1999: 9905-0132484-  
PR 04-MAY-1999: 9905-0132485-  
PR 06-MAY-1999: 9905-0132486-  
PR 07-MAY-1999: 9905-0132487-  
PR 07-MAY-1999: 9905-0132488-  
PR 11-MAY-1999: 9905-0134286-  
PR 14-MAY-1999: 9905-0134218-  
PR 14-MAY-1999: 9905-0134219-  
PR 14-MAY-1999: 9905-0134221-  
PR 14-MAY-1999: 9905-0134370-  
PR 14-MAY-1999: 9905-0134768-  
PR 14-MAY-1999: 9905-0135124-  
PR 20-MAY-1999: 9905-0135124-  
PR 21-MAY-1999: 9905-0135353-  
PR 24-MAY-1999: 9905-0135629-  
PR 25-MAY-1999: 9905-0136021-  
PR 27-MAY-1999: 9905-0136782-  
PR 28-MAY-1999: 9905-0136782-  
PR 03-JUN-1999: 9905-0137528-  
PR 04-JUN-1999: 9905-0137502-  
PR 07-JUN-1999: 9905-0137724-  
PR 08-JUN-1999: 9905-0138094-  
PR 10-JUN-1999: 9905-0138540-  
PR 10-JUN-1999: 9905-0138687-  
PR 10-JUN-1999: 9905-0138687-  
PR 16-JUN-1999: 9905-0139419-  
PR 16-JUN-1999: 9905-0139453-  
PR 17-JUN-1999: 9905-0139492-  
PR 18-JUN-1999: 9905-0139454-  
PR 18-JUN-1999: 9905-0139455-  
PR 18-JUN-1999: 9905-0139456-  
PR 18-JUN-1999: 9905-0139457-  
PR 18-JUN-1999: 9905-0139458-  
PR 18-JUN-1999: 9905-0139459-  
PR 18-JUN-1999: 9905-0139460-  
PR 18-JUN-1999: 9905-0139461-  
PR 18-JUN-1999: 9905-0139462-  
PR 18-JUN-1999: 9905-0139463-  
PR 18-JUN-1999: 9905-0139750-  
PR 18-JUN-1999: 9905-0139817-  
PR 21-JUN-1999: 9905-0139817-  
PR 22-JUN-1999: 9905-0139899-  
PR 23-JUN-1999: 9905-0140353-  
PR 23-JUN-1999: 9905-0140354-  
PR 24-JUN-1999: 9905-0140655-  
PR 24-JUN-1999: 9905-0140823-  
PR 28-JUN-1999: 9905-0140921-  
PR 28-JUN-1999: 9905-0140921-  
PR 28-JUN-1999: 9905-0141842-  
PR 01-JUL-1999: 9905-0142154-  
PR 02-JUL-1999: 9905-0142055-  
PR 06-JUL-1999: 9905-0142390-  
PR 09-JUL-1999: 9905-0142803-  
PR 09-JUL-1999: 9905-0142920-  
PR 12-JUL-1999: 9905-0143474-  
PR 13-JUL-1999: 9905-0143542-  
PR 14-JUL-1999: 9905-0143524-  
PR 15-JUL-1999: 9905-0144005-  
PR 15-JUL-1999: 9905-0144085-  
PR 16-JUL-1999: 9905-0144086-  
PR 19-JUL-1999: 9905-0144335-  
PR 19-JUL-1999: 9905-0144331-  
PR 19-JUL-1999: 9905-0144331-  
PR 19-JUL-1999: 9905-0144331-  
PR 19-JUL-1999: 9905-0144335-  
PR 20-JUL-1999: 9905-0144352-  
PR 20-JUL-1999: 9905-0144632-  
PR 20-JUL-1999: 9905-0144884-  
PR 21-JUL-1999: 9905-0144814-  
  
PR 21-JUL-1999: 9905-0145086-  
PR 21-JUL-1999: 9905-0145088-  
PR 22-JUL-1999: 9905-0145085-  
PR 22-JUL-1999: 9905-0145087-  
PR 22-JUL-1999: 9905-0145089-  
PR 22-JUL-1999: 9905-0145090-  
PR 23-JUL-1999: 9905-0145145-  
PR 23-JUL-1999: 9905-0145218-  
PR 23-JUL-1999: 9905-0145224-  
PR 26-JUL-1999: 9905-0145276-  
PR 27-JUL-1999: 9905-0145313-  
PR 27-JUL-1999: 9905-0145318-  
PR 28-JUL-1999: 9905-0145318-  
PR 28-JUL-1999: 9905-0145321-  
PR 02-AUG-1999: 9905-0146286-  
PR 02-AUG-1999: 9905-0146286-  
PR 02-AUG-1999: 9905-0146289-  
PR 03-AUG-1999: 9905-0146289-  
PR 03-AUG-1999: 9905-0147038-  
PR 04-AUG-1999: 9905-0147204-  
PR 04-AUG-1999: 9905-0147202-  
PR 05-AUG-1999: 9905-0147260-  
PR 06-AUG-1999: 9905-0147303-  
PR 06-AUG-1999: 9905-0147416-  
PR 09-AUG-1999: 9905-0147493-  
PR 09-AUG-1999: 9905-0147935-  
PR 10-AUG-1999: 9905-0148171-  
PR 11-AUG-1999: 9905-0148171-  
PR 11-AUG-1999: 9905-0148343-  
PR 13-AUG-1999: 9905-0148565-  
PR 13-AUG-1999: 9905-0148565-  
PR 16-AUG-1999: 9905-0149368-  
PR 17-AUG-1999: 9905-0149175-  
PR 18-AUG-1999: 9905-0149426-  
PR 20-AUG-1999: 9905-0149372-  
PR 20-AUG-1999: 9905-0149372-  
PR 20-AUG-1999: 9905-0149392-  
PR 23-AUG-1999: 9905-0149902-  
PR 23-AUG-1999: 9905-0149930-  
PR 25-AUG-1999: 9905-0150566-  
PR 25-AUG-1999: 9905-0150884-  
PR 27-AUG-1999: 9905-0151065-  
PR 27-AUG-1999: 9905-0151065-  
PR 27-AUG-1999: 9905-0151080-  
PR 30-AUG-1999: 9905-0151303-  
PR 31-AUG-1999: 9905-0151438-  
PR 01-SEP-1999: 9905-0151930-  
PR 07-SEP-1999: 9905-0152363-  
PR 10-SEP-1999: 9905-0153070-  
PR 11-SEP-1999: 9905-0153078-  
PR 11-SEP-1999: 9905-0153078-  
PR 15-SEP-1999: 9905-0154039-  
PR 16-SEP-1999: 9905-0154039-  
PR 20-SEP-1999: 9905-0154779-  
PR 22-SEP-1999: 9905-0155139-  
PR 23-SEP-1999: 9905-0155486-  
PR 24-SEP-1999: 9905-0155659-  
PR 28-SEP-1999: 9905-0156248-  
PR 28-SEP-1999: 9905-0156349-  
PR 04-OCT-1999: 9905-0157117-  
PR 05-OCT-1999: 9905-0157273-  
PR 06-OCT-1999: 9905-0157865-  
PR 06-OCT-1999: 9905-0158029-  
PR 08-OCT-1999: 9905-0158222-  
PR 12-OCT-1999: 9905-0158369-  
PR 12-OCT-1999: 9905-0159233-  
PR 13-OCT-1999: 9905-0159233-  
PR 13-OCT-1999: 9905-0159235-  
PR 14-OCT-1999: 9905-0159330-  
PR 14-OCT-1999: 9905-0159331-  
PR 14-OCT-1999: 9905-0159637-  
PR 14-OCT-1999: 9905-0159638-  
PR 16-OCT-1999: 9905-0159584-

PR 21-OCT-1999; 990S-0160741.  
 PR 21-OCT-1999; 990S-0160750.  
 PR 21-OCT-1999; 990S-0160760.  
 PR 21-OCT-1999; 990S-0160770.  
 PR 21-OCT-1999; 990S-0160814.  
 PR 21-OCT-1999; 990S-0160815.  
 PR 22-OCT-1999; 990S-0160980.  
 PR 22-OCT-1999; 990S-0160981.  
 PR 22-OCT-1999; 990S-0160989.  
 PR 22-OCT-1999; 990S-0161404.  
 PR 22-OCT-1999; 990S-0161406.  
 PR 25-OCT-1999; 990S-0161359.  
 PR 26-OCT-1999; 990S-0161360.  
 PR 26-OCT-1999; 990S-0161361.  
 PR 26-OCT-1999; 990S-0161920.  
 PR 28-OCT-1999; 990S-0161992.  
 PR 28-OCT-1999; 990S-0161993.  
 PR 28-OCT-1999; 990S-0162142.

Query Match 100.0%; Score 33; DB 21; Length 552;  
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSC 6  
 DB 78 999199 83

RESULT 27

AAB94160  
 ID AAB94160 standard; Protein: 557 AA.

XX AAB94160.  
 XX

DT 26-JUN-2001 (first entry)  
 XX

DE Human protein sequence SEQ ID NO:14452.  
 XX

KM Human: primer: detection; diagnosis; antisense therapy; gene therapy.  
 XX

OK Homo sapiens.  
 XX

PN EP1074617-A2.  
 XX

PD 07-FEB-2001.  
 XX

PF 28-JUL-2000; 2000EP-0116126.  
 XX

PK 29-JUL-1999; 99JP-0248036.  
 XX

PR 27-AUG-1999; 99JP-01300253.  
 XX

PR 11-JAN-2000; 2000JP-018776.  
 XX

PR 02-MAY-2000; 2000JP-0183767.  
 XX

PR 09-JUN-2000; 2000JP-0241899.  
 XX

XX (HELI-) HELIX RES INST.  
 XX

PI Ora T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX

DR WPI: 2001-318749/34.  
 XX

PR Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT full-length cDNAs of the abnormality of the proteins encoded by the  
 PT full-length cDNAs  
 XX

PS Claim 8: SEQ ID 14452; 2537pp + CD ROM; English.  
 XX

CC The present invention describes primer sets for synthesising 5602  
 CC full-length cDNAs defined in the specification. Where a primer set  
 CC full-length cDNAs defined in the specification, and an oligonucleotide complementary  
 CC to the complementary strand of a polynucleotide which comprises one of

CC the 5602 nucleotide sequences defined in the specification, where the  
 CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 CC of an oligonucleotide comprising a sequence complementary to the  
 CC sequence and an oligonucleotide comprising a sequence complementary to a  
 CC polynucleotide which comprises a 3'-end sequence, where the  
 CC oligonucleotide comprises at least 15 nucleotides and the combination of  
 CC the 5'-end sequence/3'-end sequence is selected from those defined in  
 CC the specification. The primer sets can be used in antisense therapy and  
 CC in gene therapy. The primers are useful for synthesizing polynucleotides,  
 CC particularly full-length cDNAs. The primers are also useful for the  
 CC detection of the abnormality of the proteins encoded by  
 CC the full-length cDNAs. The primers are also useful for the detection of  
 CC cDNAs easily without any specialised methods. AAB94160 to AAB13628 and  
 CC AAB13633 represent human amino acid sequences; AAB92446 to  
 CC AAB13632 represent human amino acid sequences; and AAB13632  
 CC represent oligonucleotides, all of which are used in the exemplification  
 CC of the present invention.  
 XX

Sequence 557 AA:

Query Match 100.0%; Score 33; DB 22; Length 557;  
 Best Local Similarity 100.0%; Pred. No. 7.5e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSC 6  
 DB 431 999199 436

RESULT 28

AAG81993  
 ID AAG81993 standard; Protein: 582 AA.

XX AAG81993.  
 XX

DT 03-SEP-2001 (first entry)  
 XX

DE S. epidermidis open reading frame protein sequence SEQ ID NO:1080.  
 XX

KM Staphylococcus epidermidis SRI strain; Infection; diagnosis;  
 XX vaccination; endocarditis.  
 XX

OK Staphylococcus epidermidis.  
 XX

PN WO200134809-A2.  
 XX

PD 17-MAY-2001.  
 XX

PK 09-NOV-2000; 2000MO-US07872.  
 XX

PR 09-NOV-1999; 99US-0164258.  
 XX

XX (GLAX) GLAXO GROUP LTD.  
 XX

PA Kimmerly MJ;  
 XX  
 PT WPI: 2001-316495/33.  
 XX

DR N-PSDB: AAB52843.  
 XX

PR Nucleic acids encoding polypeptides from Staphylococcus epidermidis,  
 PT useful for vaccinating against infections, e.g. endocarditis -  
 PT  
 XX

PS Claim 18: Page 315; 2188pp; English.  
 XX

CC AAB5304 to AAB53970 represent nucleic acids (I) encoding polypeptides  
 CC (II), given in AAG81454 to AAG83120, from Staphylococcus epidermidis.  
 CC (I) and (II) can have antibacterial activity and therefore can be used  
 CC in vaccination. The nucleic acids (I) may be used to produce the  
 CC S. epidermidis polypeptides (II) via the production of vectors  
 CC containing them which are used to produce host cells which express the  
 CC polypeptides. The polypeptides (II) (and/or nucleic acids) may then be



Novel nucleic acids and polypeptides, useful for treating disorders such as central nervous system injuries -

Example 4: SEQ ID NO 3148: 10078pp; English.

The invention relates to human nucleic acids (AA157298-AA161369) and the encoded polypeptides (AA163662-AA162213) with neurotrophic immunosuppressant and cytoskeletal activity. The polynucleotides are useful in gene therapy. A composition containing a polypeptide or polynucleotide of the invention may be used to treat diseases of the peripheral nervous system, such as peripheral nervous injuries, peripheral neuropathy and Alzheimer's, Parkinson's, Huntington's disease, amyotrophic lateral sclerosis, and Shy-Drager Syndrome. The polynucleotides are useful in the activation of the activities such as: immune system suppression, activation/inhibition activity, chemotactic/chemokinetic activity, haemostatic and thrombolytic activity, cancer diagnosis and therapy, drug screening, assays for receptor activity, arthritis and inflammation, leukemias and cancer disorders.

Note: The sequence data for this patent did not form part of the printed specification.

Sequence 525 AA:

Query Match Best Local Similarity 100.0%; Score 33; DB 22; Length 625;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGTSG 6

106 999GTCG 111

RESULT 31

AA894603 ID AAB94603 standard; Protein: 678 AA.

AC AAB94603;

26-JUN-2001 (first entry)

Human protein sequence SEQ ID NO:15428.

Human; primer: detection; diagnosis; antisense therapy; gene therapy.

Homo sapiens.

EP1074617-A2.

07-FEB-2001.

28-JUL-2000; 2000EP-0116126.

29-JUL-1999; 99JP-0248036.

22-JUL-1999; 99JP-0248036.

11-JAN-2000; 2000JP-018775.

02-MAY-2000; 2000JP-018767.

09-JUN-2000; 2000JP-0241899.

(HELT-) HELIX RES INSP.

Ota T, Isegaki T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

WPI: 2001-318749/34.

Primer sets for synthesizing polynucleotides, particularly the 5602 full-length cDNAs defined in the specification, and for the detection and diagnosis of the abnormality of the proteins encoded by the full-length cDNAs -

Claim 8: SEQ ID 15428; 2537pp + CD ROM; English.

The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification, where a primer set comprises: (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 nucleotide sequences defined in the specification, where the oligonucleotide comprises at least 15 nucleotides; or (b) a combination of two oligonucleotides comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a sequence complementary to a polynucleotide which comprises a 3'-end sequence, where the oligonucleotide comprises at least 15 nucleotides and the combination of the 5'-end sequence/3'-end sequence is selected from those defined in the specification. The primer sets can be used in antisense therapy and particularly, the primers are useful for synthesizing polynucleotides, detection and/or diagnosis of the abnormality of the full-length cDNAs. The primers allow obtaining of the full-length cDNAs easily without any specialised methods. AA103166 to AA13628 and AA13632 to AA18742 represent human cDNA sequences; AAB92446 to AAB95833 represent human amino acid sequences; and AA13629 to AA13632 represent oligonucleotides, all of which are used in the exemplification of the present invention.

Sequence 678 AA:

Query Match Best Local Similarity 100.0%; Score 33; DB 22; Length 678;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GCGTSG 6

659 999GTCG 664

RESULT 32

AA184686 ID AAB184686 standard; Protein: 2071 AA.

AA184686;

08-AUG-2000 (first entry)

Amino acid sequence of AC003, a GTPase activating protein.

GTP-binding protein; AC003; GTPase activating protein; fungal growth;

phosphatidylinositol-4-kinase; cytokinesis gene; fungal development;

fungicide; filamentous fungi; plant pathogen; septicemia; tritic;.

Stagnospora nodorum; Magnaporthe oryzae; human pathogen;

Candida albicans; Aspergillus fumigatus.

Asbya gossypii.

WO2000022133-A1.

20-APR-2000.

06-OCT-1999; 99MO-EP07501.

08-OCT-1998; 98US-0168804.

(NOVS) NOVARTIS AG.

(NOVS) NOVARTIS-ERFINDUNGEN VERW GES MBH.

Gaffney TD, Wendland J, Dietrich F, Phillipsen P, Coff SA;

WPI: 2000-317988/27.

N-PSDB: AA141503.

Asbya gossypii nucleotide and protein sequences useful for identifying fungal growth inhibitors -

PS Claim 2: Page 48-60: 93pp: English.

XX The present sequence represents a putative GTPase activating protein,  
 CC does not appear to be a GTPase activating protein.  
 CC binding protein genes, putative phosphatidylinositol-4-kinase  
 CC protein gene, and a putative cyclohexis gene. These genes are  
 CC essential for fungal growth and development. The proteins can be used  
 CC in methods to identify compounds that have fungicidal activity. Compounds  
 CC with fungicidal activity can be used for fungicidal activity. Compounds  
 CC especially of filamentous fungi, fungi that can be suppressed include  
 CC *Aspergillus fumigatus*, *Aspergillus niger*, *Aspergillus nidulans*, and  
 CC *Marasmius oreades*, and human pathogens (e.g., candida albicans, and  
 CC *Aspergillus fumigatus*).

CC Sequence 2071 AA:

QY 1 GCGTSG 6  
 DB 199 99GTSG 204  
 111111

Query Match 100.0%: Score 33: DB 21: Length 2071;  
 Best Local Similarity 100.0%: Pred. No. 2.4e+03;  
 Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

RESULT 33

AA86196  
 ID AAB6196 standard: Protein: 3025 AA.  
 AC AAB6196:  
 DE 21-MUG-2001 (first entry)  
 DE HIV-1 subtype C protein fragment #2.  
 DE HIV-1 subtype C protein fragment #2.  
 KM Infection: diagnosis: human: humoral immune response: antiviral:  
 KM cellular immune response: vaccine: treatment: gene therapy.  
 OS human immunodeficiency virus type 1.  
 PA Key location/Qualifiers  
 FT Misc-difference 1..3025  
 FT //label= Xaa  
 FT /note= "Xaa represents a stop codon"  
 XX De10056747-A1.  
 XX 31-MAY-2001.  
 XX 16-NOV-2000: 2000DE-1056747.  
 XX 16-NOV-1999: 99DE-1055089.  
 XX (SHAO) SHAO Y  
 XX (GENE) GENENT GSH GGS ANGENANDE BIOTECHNOLOG.  
 XX Wagner R, Wolf H, Shao Y, Graf M:  
 PI WPI: 2001-336417/36.  
 DR New nucleic acid sequences from a human immune deficiency virus  
 PT Intersubtype, useful for treatment, prevention and diagnosis of  
 PT infection -  
 CC Disclosure: Fig 8A-O: 48pp: German.

XX This invention describes a novel polynucleotide isolated from human  
 CC immunodeficiency virus type 1 subtype C/B' which can be used for the  
 CC induction of specific humoral and cellular immune responses. (1) and  
 CC polypeptides (11) encoded by them, are useful in pharmaceuticals,  
 CC particularly in the treatment of HIV-1 infections, also for rational  
 CC of human immune deficiency virus 1 (HIV-1) infections, also for rational

CC design of test or therapeutic reagents, or gene therapy vectors.  
 CC Polypeptides, especially antibodies, specifically directed against (11)  
 CC are similarly useful as pharmaceutical and diagnostic agents. (1) are  
 CC particularly useful in the treatment of HIV-1 infections. (1) are  
 CC (particularly China and South-East Asia) where this subtype is prevalent.  
 CC The products of the invention have antiviral activity. This sequence  
 CC represents a protein encoded by the HIV-1 subtype C genome described in  
 CC the method of the invention.

CC Sequence 3025 AA:

QY 1 GCGTSG 6  
 DB 2639 99GTSG 2644  
 111111

Query Match 100.0%: Score 33: DB 22: Length 3025;  
 Best Local Similarity 100.0%: Pred. No. 3.4e+03;  
 Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

RESULT 34

AA890042  
 ID AAM90042 standard: Protein: 32 AA.  
 AC AAM90042:  
 DE 26-PEB-1999 (first entry)  
 DE B. stearotherophilus alanine dehydrogenase fragment #1.  
 DE B. stearotherophilus alanine dehydrogenase fragment #1.  
 KM Phenylalanine dehydrogenase: pdh: L-amino acid: phenylpyruvic acid:  
 KM aspartame production: phenylketonuria: infant: alanine dehydrogenase:  
 KM diet.  
 OS Bacillus stearotherophilus.  
 PA Bacillus stearotherophilus.  
 FT US8581810-A.  
 PD 22-DEC-1998.  
 XX 05-JUN-1995: 95US-0461990.  
 XX 05-JUN-1995: 95US-0461990.  
 XX (YESH ) UNIV YESHIVA EINSTEIN COLLEGE.  
 PA Blanchard JS:  
 PI WPI: 1999-080406/07.  
 XX DNA encoding Rhodococcus L-phenylalanine dehydrogenase - for  
 PT production of recombinant enzyme  
 CC disclosure: Fig 4: 30pp: English.

XX This sequence is used to describe a method which results in the isolation  
 CC of a phenylalanine dehydrogenase protein, pdh, from Rhodococcus sp. M4.  
 CC The encoding pdh nucleic acid is used to produce recombinant Rhodococcus  
 CC L-phenylalanine dehydrogenase, which can be used to produce L-amino acids  
 CC from the corresponding alpha-keto acids, e.g. to produce L-phenylalanine  
 CC from phenylpyruvic acid (e.g. for the production of aspartame), to screen  
 CC for phenylketonuria in newborn infants and to monitor the efficacy of  
 CC low-phenylalanine diets.

XX Sequence 32 AA:

QY 1 GCGTSG 6  
 DB 199 99GTSG 204  
 111111

Query Match 90.9%: Score 30: DB 20: Length 32;  
 Best Local Similarity 83.3%: Pred. No. 1.8e+02;  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:



DB 6 999tag 11

RESULT 35

AAW77518 standard; protein: 66 AA.

AAW77518:

30-OCT-1998 (first entry)

Staphylococcus aureus protein of unknown function.

Staphylococcus aureus protein; immune response induction; eye infection; antibody production; T-cell immune response; gastrointestinal infection; respiratory infection; inhibitor; bacterial infection; cardiac infection; respiratory virus system; kidney infection; urinary tract infection; antimicrobial compound identification; broad spectrum antibiotic; therapy.

Staphylococcus aureus.

Key location/Qualifiers

Misc-difference 1:86 \*residues designated X are unspecified, and /note represented as X in the specification\*

EP841394-A2.

13-MAY-1998.

24-SEP-1997: 97EP-0107485.

24-SEP-1996: 96US-0027032.

(SMK) SMITHKLINE BEECHAM CORP.

(SMK) SMITHKLINE BEECHAM PLC.

Black Wf, Burham MKR, Hodgson JB, Knowles DTC, Lomanto MA, Nicholas RO, Ratte JM, Reichard RW, Rosenberg M, Ward JW;

WPI: 1998-252940/23.

N-PSDB: AAV53319.

New nucleic acid sequences from Staphylococcus aureus WCH029 - useful in vaccine development

respiratory tract and central nervous system

Claim 11: Page 228: 390pp: English.

This sequence represents a Staphylococcus aureus protein of unknown function, and is encoded by a DNA sequence of the invention.

The protein was isolated from Staphylococcus aureus WCH029 (NCIMB 40771). Host cells were transfected with the sequence used to produce polypeptides or fragments. The proteins were used in treatment of disease, for inducing an immune response by administering them to produce antibody and/or T-cell immune response. Antagonists of the proteins are used for the inhibition of bacterial infections, especially conditions which may be treated include bacterial infections, especially urinary tract infection, gastroenteric, central nervous, eye, kidney, to identify antimicrobial compounds which can be used in combination with especially antimicrobial compounds in the treatment of H. pylori infection.

Sequence 66 AA:

Query Match 90.94; Score 30; DB 19; Length 66;

Best Local Similarity 83.34; Pred No 34e102;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

0Y 1 GCGTSG 6

DB 14 999tag 19

RESULT 36

AA015636 standard; protein: 81 AA.

AA015636:

17-OCT-2000 (first entry)

Arabidopsis thaliana protein fragment SEQ ID NO: 15966.

Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; terminator sequence.

Arabidopsis thaliana.

EP103405-A2.

06-SEP-2000.

25-FEB-2000: 2000EP-0301439.

25-FEB-1999: 99US-0121825.

05-MAR-1999: 99US-0121830.

05-MAR-1999: 99US-0123548.

23-MAR-1999: 99US-0125788.

25-MAR-1999: 99US-0126264.

29-MAR-1999: 99US-0126785.

01-APR-1999: 99US-0127462.

06-APR-1999: 99US-0128234.

13-APR-1999: 99US-0128409.

16-APR-1999: 99US-0128645.

19-APR-1999: 99US-0130077.

21-APR-1999: 99US-0130449.

23-APR-1999: 99US-0130510.

23-APR-1999: 99US-0130891.

28-APR-1999: 99US-0131449.

30-APR-1999: 99US-0132009.

04-MAY-1999: 99US-0132484.

05-MAY-1999: 99US-0132485.

06-MAY-1999: 99US-0132486.

06-MAY-1999: 99US-0132487.

07-MAY-1999: 99US-0132863.

14-MAY-1999: 99US-0134256.

14-MAY-1999: 99US-0134257.

14-MAY-1999: 99US-0134261.

14-MAY-1999: 99US-0134271.

18-MAY-1999: 99US-0134370.

18-MAY-1999: 99US-0134768.

19-MAY-1999: 99US-0134941.

20-MAY-1999: 99US-0135124.

21-MAY-1999: 99US-0135233.

24-MAY-1999: 99US-0136021.

27-MAY-1999: 99US-0136392.

28-MAY-1999: 99US-0136782.

01-JUN-1999: 99US-0137222.

03-JUN-1999: 99US-0137528.

07-JUN-1999: 99US-0137532.

07-JUN-1999: 99US-0137714.

08-JUN-1999: 99US-0138054.

10-JUN-1999: 99US-0138540.

10-JUN-1999: 99US-0138847.

14-JUN-1999: 99US-0139119.

16-JUN-1999: 99US-0139452.

15-JUN-1999: 99US-0139453.

17-JUN-1999: 99US-0139492.

PR 18-JUN-1999: 9905-0139454.  
 PR 18-JUN-1999: 9905-0139455.  
 PR 18-JUN-1999: 9905-0139456.  
 PR 18-JUN-1999: 9905-0139457.  
 PR 18-JUN-1999: 9905-0139458.  
 PR 18-JUN-1999: 9905-0139459.  
 PR 18-JUN-1999: 9905-0139460.  
 PR 18-JUN-1999: 9905-0139461.  
 PR 18-JUN-1999: 9905-0139462.  
 PR 18-JUN-1999: 9905-0139463.  
 PR 18-JUN-1999: 9905-0139464.  
 PR 18-JUN-1999: 9905-0139465.  
 PR 22-JUN-1999: 9905-0138899.  
 PR 23-JUN-1999: 9905-0140353.  
 PR 23-JUN-1999: 9905-0140354.  
 PR 24-JUN-1999: 9905-0140355.  
 PR 24-JUN-1999: 9905-0140356.  
 PR 28-JUN-1999: 9905-0140823.  
 PR 28-JUN-1999: 9905-0140921.  
 PR 28-JUN-1999: 9905-0140922.  
 PR 01-JUL-1999: 9905-0141187.  
 PR 02-JUL-1999: 9905-0142154.  
 PR 06-JUL-1999: 9905-0142055.  
 PR 08-JUL-1999: 9905-0142390.  
 PR 09-JUL-1999: 9905-0142603.  
 PR 11-JUL-1999: 9905-0142777.  
 PR 11-JUL-1999: 9905-0142778.  
 PR 14-JUL-1999: 9905-0143624.  
 PR 15-JUL-1999: 9905-0144005.  
 PR 16-JUL-1999: 9905-0144085.  
 PR 16-JUL-1999: 9905-0144086.  
 PR 19-JUL-1999: 9905-0144325.  
 PR 19-JUL-1999: 9905-0144331.  
 PR 19-JUL-1999: 9905-0144332.  
 PR 19-JUL-1999: 9905-0144333.  
 PR 19-JUL-1999: 9905-0144334.  
 PR 19-JUL-1999: 9905-0144335.  
 PR 20-JUL-1999: 9905-0144352.  
 PR 20-JUL-1999: 9905-0144682.  
 PR 20-JUL-1999: 9905-0144683.  
 PR 21-JUL-1999: 9905-0145086.  
 PR 21-JUL-1999: 9905-0145087.  
 PR 22-JUL-1999: 9905-0145085.  
 PR 22-JUL-1999: 9905-0145087.  
 PR 22-JUL-1999: 9905-0145089.  
 PR 22-JUL-1999: 9905-0145192.  
 PR 22-JUL-1999: 9905-0145218.  
 PR 23-JUL-1999: 9905-0145224.  
 PR 26-JUL-1999: 9905-0145276.  
 PR 27-JUL-1999: 9905-0145913.  
 PR 27-JUL-1999: 9905-0145918.  
 PR 27-JUL-1999: 9905-0145919.  
 PR 28-JUL-1999: 9905-0145931.  
 PR 28-JUL-1999: 9905-0145931.  
 PR 02-AUG-1999: 9905-0146388.  
 PR 02-AUG-1999: 9905-0146389.  
 PR 03-AUG-1999: 9905-0147036.  
 PR 04-AUG-1999: 9905-0147204.  
 PR 04-AUG-1999: 9905-0147302.  
 PR 05-AUG-1999: 9905-0147122.  
 PR 05-AUG-1999: 9905-0147303.  
 PR 06-AUG-1999: 9905-0147416.  
 PR 09-AUG-1999: 9905-0147493.  
 PR 09-AUG-1999: 9905-0147935.  
 PR 10-AUG-1999: 9905-0148171.  
 PR 11-AUG-1999: 9905-0148319.  
 PR 13-AUG-1999: 9905-0148341.  
 PR 13-AUG-1999: 9905-0148684.

PR 16-AUG-1999: 9905-0149368.  
 PR 17-AUG-1999: 9905-0149175.  
 PR 18-AUG-1999: 9905-0149426.  
 PR 20-AUG-1999: 9905-0149722.  
 PR 20-AUG-1999: 9905-0149723.  
 PR 20-AUG-1999: 9905-0149920.  
 PR 23-AUG-1999: 9905-0149902.  
 PR 23-AUG-1999: 9905-0149930.  
 PR 25-AUG-1999: 9905-0150566.  
 PR 26-AUG-1999: 9905-0150884.  
 PR 27-AUG-1999: 9905-0151065.  
 PR 27-AUG-1999: 9905-0151066.  
 PR 30-AUG-1999: 9905-0151080.  
 PR 31-AUG-1999: 9905-0151303.  
 PR 01-SEP-1999: 9905-0151438.  
 PR 10-SEP-1999: 9905-0152363.  
 PR 10-SEP-1999: 9905-0153070.  
 PR 11-SEP-1999: 9905-0153159.  
 PR 11-SEP-1999: 9905-0153160.  
 PR 16-SEP-1999: 9905-0154039.  
 PR 20-SEP-1999: 9905-0154779.  
 PR 22-SEP-1999: 9905-0155133.  
 PR 23-SEP-1999: 9905-0155486.  
 PR 24-SEP-1999: 9905-0155559.  
 PR 28-SEP-1999: 9905-0156438.  
 PR 28-SEP-1999: 9905-0156439.  
 PR 01-OCT-1999: 9905-0157117.  
 PR 05-OCT-1999: 9905-0157753.  
 PR 06-OCT-1999: 9905-0157865.  
 PR 07-OCT-1999: 9905-0158029.  
 PR 08-OCT-1999: 9905-0158232.  
 PR 12-OCT-1999: 9905-0158369.  
 PR 12-OCT-1999: 9905-0158370.  
 PR 13-OCT-1999: 9905-0158392.  
 PR 13-OCT-1999: 9905-0159299.  
 PR 14-OCT-1999: 9905-0159329.  
 PR 14-OCT-1999: 9905-0159330.  
 PR 14-OCT-1999: 9905-0159331.  
 PR 14-OCT-1999: 9905-0159337.  
 PR 14-OCT-1999: 9905-0159338.  
 PR 14-OCT-1999: 9905-0159340.  
 PR 14-OCT-1999: 9905-0159341.  
 PR 21-OCT-1999: 9905-0160741.  
 PR 21-OCT-1999: 9905-0160747.  
 PR 21-OCT-1999: 9905-0160768.  
 PR 21-OCT-1999: 9905-0160770.  
 PR 21-OCT-1999: 9905-0160814.  
 PR 21-OCT-1999: 9905-0160815.  
 PR 21-OCT-1999: 9905-0160816.  
 PR 21-OCT-1999: 9905-0160817.  
 PR 22-OCT-1999: 9905-0160981.  
 PR 22-OCT-1999: 9905-0160989.  
 PR 25-OCT-1999: 9905-0161404.  
 PR 25-OCT-1999: 9905-0161405.  
 PR 25-OCT-1999: 9905-0161406.  
 PR 26-OCT-1999: 9905-0161359.  
 PR 26-OCT-1999: 9905-0161360.  
 PR 26-OCT-1999: 9905-0161361.  
 PR 28-OCT-1999: 9905-0161920.  
 PR 28-OCT-1999: 9905-0161992.  
 PR 28-OCT-1999: 9905-0161993.  
 PR 29-OCT-1999: 9905-0162142.

Query Match 90.98: Score 30, PB 21, Length 81:  
 1: 100% Identical Similarity 83.3%: Tree No. 46002, 0: Indels 0: Caps 0:  
 Matches 5: Conservative 1: Mismatches 0:

QY 1 GGCTSG 6  
 Db 35 ggcgtc 40

RESULT 37  
 ANG15635

ID AAG15635 standard; Protein: 83 AA.  
XX AAG15635;  
AC  
XX  
XX 17-OCT-2000 (first entry)  
DE Arabidopsis thaliana protein fragment SEQ ID NO: 15965.  
XX  
XX  
XX Protein identification: signal transduction pathway; metabolic pathway;  
XX hydrolysis assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
OS Arabidopsis thaliana.  
PM EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000: 2000EP-020439.  
XX  
PR 25-FEB-1999: 9905-012825.  
PR 05-MAR-1999: 9905-0123180.  
PR 09-MAR-1999: 9905-0123548.  
PR 23-MAR-1999: 9905-0125788.  
PR 23-MAR-1999: 9905-0128254.  
PR 23-MAR-1999: 9905-0127482.  
PR 01-APR-1999: 9905-0128234.  
PR 06-APR-1999: 9905-0128714.  
PR 08-APR-1999: 9905-0128845.  
PR 16-APR-1999: 9905-0130077.  
PR 19-APR-1999: 9905-0130449.  
PR 21-APR-1999: 9905-0130530.  
PR 23-APR-1999: 9905-0131449.  
PR 28-APR-1999: 9905-0132048.  
PR 30-APR-1999: 9905-0132407.  
PR 04-MAY-1999: 9905-0132485.  
PR 05-MAY-1999: 9905-0132486.  
PR 08-MAY-1999: 9905-0132487.  
PR 07-MAY-1999: 9905-0132486.  
PR 11-MAY-1999: 9905-0134256.  
PR 14-MAY-1999: 9905-0134218.  
PR 14-MAY-1999: 9905-0134219.  
PR 14-MAY-1999: 9905-0134221.  
PR 14-MAY-1999: 9905-0134370.  
PR 19-MAY-1999: 9905-0134708.  
PR 20-MAY-1999: 9905-0135124.  
PR 21-MAY-1999: 9905-0135353.  
PR 24-MAY-1999: 9905-0135629.  
PR 25-MAY-1999: 9905-0136021.  
PR 27-MAY-1999: 9905-0136392.  
PR 29-MAY-1999: 9905-0136752.  
PR 01-JUN-1999: 9905-0137487.  
PR 03-JUN-1999: 9905-0137502.  
PR 04-JUN-1999: 9905-0137502.  
PR 07-JUN-1999: 9905-0137724.  
PR 08-JUN-1999: 9905-0138094.  
PR 10-JUN-1999: 9905-0138540.  
PR 10-JUN-1999: 9905-0138647.  
PR 14-JUN-1999: 9905-0139452.  
PR 16-JUN-1999: 9905-0139453.  
PR 17-JUN-1999: 9905-0139452.  
PR 18-JUN-1999: 9905-0139454.  
PR 18-JUN-1999: 9905-0139455.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139460.

PR 18-JUN-1999: 9905-0139463.  
PR 18-JUN-1999: 9905-0139462.  
PR 18-JUN-1999: 9905-0139463.  
PR 18-JUN-1999: 9905-0139750.  
PR 18-JUN-1999: 9905-0139763.  
PR 21-JUN-1999: 9905-0139817.  
PR 22-JUN-1999: 9905-0139899.  
PR 23-JUN-1999: 9905-0140353.  
PR 23-JUN-1999: 9905-0140353.  
PR 24-JUN-1999: 9905-0140655.  
PR 28-JUN-1999: 9905-0140833.  
PR 29-JUN-1999: 9905-0140991.  
PR 30-JUN-1999: 9905-0141287.  
PR 01-JUL-1999: 9905-0141842.  
PR 02-JUL-1999: 9905-0142154.  
PR 06-JUL-1999: 9905-0142390.  
PR 08-JUL-1999: 9905-0142803.  
PR 09-JUL-1999: 9905-0142920.  
PR 12-JUL-1999: 9905-0142977.  
PR 13-JUL-1999: 9905-0143542.  
PR 14-JUL-1999: 9905-0143542.  
PR 15-JUL-1999: 9905-0143604.  
PR 16-JUL-1999: 9905-0144085.  
PR 16-JUL-1999: 9905-0144086.  
PR 19-JUL-1999: 9905-0144335.  
PR 19-JUL-1999: 9905-0144331.  
PR 19-JUL-1999: 9905-0144332.  
PR 19-JUL-1999: 9905-0144333.  
PR 19-JUL-1999: 9905-0144335.  
PR 20-JUL-1999: 9905-0144332.  
PR 20-JUL-1999: 9905-0144632.  
PR 20-JUL-1999: 9905-0144684.  
PR 21-JUL-1999: 9905-0144814.  
PR 21-JUL-1999: 9905-0145086.  
PR 22-JUL-1999: 9905-0145088.  
PR 22-JUL-1999: 9905-0145087.  
PR 22-JUL-1999: 9905-0145087.  
PR 22-JUL-1999: 9905-0145182.  
PR 23-JUL-1999: 9905-0145185.  
PR 23-JUL-1999: 9905-0145218.  
PR 23-JUL-1999: 9905-0145224.  
PR 23-JUL-1999: 9905-0145224.  
PR 27-JUL-1999: 9905-0145315.  
PR 27-JUL-1999: 9905-0145318.  
PR 27-JUL-1999: 9905-0145318.  
PR 28-JUL-1999: 9905-0145391.  
PR 28-JUL-1999: 9905-0145386.  
PR 02-AUG-1999: 9905-0146388.  
PR 02-AUG-1999: 9905-0147038.  
PR 03-AUG-1999: 9905-0147038.  
PR 04-AUG-1999: 9905-0147204.  
PR 04-AUG-1999: 9905-0147302.  
PR 05-AUG-1999: 9905-0147132.  
PR 05-AUG-1999: 9905-0147280.  
PR 06-AUG-1999: 9905-0147303.  
PR 06-AUG-1999: 9905-0147445.  
PR 09-AUG-1999: 9905-0147955.  
PR 10-AUG-1999: 9905-0148111.  
PR 11-AUG-1999: 9905-0148319.  
PR 11-AUG-1999: 9905-0148319.  
PR 12-AUG-1999: 9905-0148365.  
PR 13-AUG-1999: 9905-0148365.  
PR 16-AUG-1999: 9905-0148968.  
PR 17-AUG-1999: 9905-0149375.  
PR 18-AUG-1999: 9905-0149426.  
PR 20-AUG-1999: 9905-0149723.  
PR 20-AUG-1999: 9905-0149723.  
PR 25-AUG-1999: 9905-0149929.  
PR 25-AUG-1999: 9905-0149929.

PR 23-AUG-1999: 99US-0149300.  
 PR 25-AUG-1999: 99US-0150566.  
 PR 26-AUG-1999: 99US-0150884.  
 PR 27-AUG-1999: 99US-0151065.  
 PR 27-AUG-1999: 99US-0151066.  
 PR 27-AUG-1999: 99US-0151067.  
 PR 27-AUG-1999: 99US-0151068.  
 PR 30-AUG-1999: 99US-0151300.  
 PR 31-AUG-1999: 99US-0151438.  
 PR 01-SEP-1999: 99US-0151930.  
 PR 07-SEP-1999: 99US-0152363.  
 PR 10-SEP-1999: 99US-0153070.  
 PR 12-SEP-1999: 99US-0153758.  
 PR 13-SEP-1999: 99US-0153759.  
 PR 15-SEP-1999: 99US-0154038.  
 PR 20-SEP-1999: 99US-0154779.  
 PR 22-SEP-1999: 99US-0155139.  
 PR 23-SEP-1999: 99US-0155486.  
 PR 24-SEP-1999: 99US-0155659.  
 PR 26-SEP-1999: 99US-0156458.  
 PR 27-SEP-1999: 99US-0156459.  
 PR 28-SEP-1999: 99US-0157217.  
 PR 05-OCT-1999: 99US-0157753.  
 PR 06-OCT-1999: 99US-0157865.  
 PR 07-OCT-1999: 99US-0158029.  
 PR 08-OCT-1999: 99US-0158232.  
 PR 12-OCT-1999: 99US-0158369.  
 PR 13-OCT-1999: 99US-0159283.  
 PR 14-OCT-1999: 99US-0159284.  
 PR 14-OCT-1999: 99US-0159285.  
 PR 14-OCT-1999: 99US-0159329.  
 PR 14-OCT-1999: 99US-0159330.  
 PR 14-OCT-1999: 99US-0159331.  
 PR 14-OCT-1999: 99US-0159637.  
 PR 14-OCT-1999: 99US-0159638.  
 PR 14-OCT-1999: 99US-0160741.  
 PR 21-OCT-1999: 99US-0160767.  
 PR 21-OCT-1999: 99US-0160768.  
 PR 21-OCT-1999: 99US-0160770.  
 PR 21-OCT-1999: 99US-0160814.  
 PR 21-OCT-1999: 99US-0160815.  
 PR 21-OCT-1999: 99US-0160820.  
 PR 22-OCT-1999: 99US-0160880.  
 PR 22-OCT-1999: 99US-0160889.  
 PR 25-OCT-1999: 99US-0161404.  
 PR 25-OCT-1999: 99US-0161405.  
 PR 25-OCT-1999: 99US-0161406.  
 PR 26-OCT-1999: 99US-0161359.  
 PR 26-OCT-1999: 99US-0161360.  
 PR 28-OCT-1999: 99US-0161920.  
 PR 28-OCT-1999: 99US-0161992.  
 PR 28-OCT-1999: 99US-0161993.  
 PR 29-OCT-1999: 99US-0162142.

Query Match 90.9%; Score 30; DB 21; Length 83;  
 Best Local Similarity 83.3%; Pred. No. 4, e+02;  
 Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTSG 6  
 Db 37 g99ftg 42

RESULT 38  
 ID AAW90054 standard; Protein: 95 AA.  
 AC AAW90054:  
 XX 26-FEB-1999 (first entry)  
 DT 26-FEB-1999 (first entry)  
 XX B. steaerothermophilus alanine dehydrogenase fragment #2.  
 DE

XX Phenylalanine dehydrogenase; pdh; L-amino acid; phenylpyruvic acid;  
 KW aspartame production; phenylketonuria; infant; alanine dehydrogenase;  
 KW diet.  
 XX  
 XX Bacillus steaerothermophilus.  
 XX  
 XX US5851810-A.  
 PN  
 PD 22-DEC-1998.  
 XX  
 XX 05-JUN-1995: 99US-0461990.  
 XX  
 XX 05-JUN-1995: 99US-0461990.  
 XX  
 PA (YESH ) UNIV YESHIVA EINSTEIN COLLEGE.  
 XX  
 XX Blanchard JS;  
 PI  
 XX WPI: 1999-080406/07.  
 XX  
 PT DNA encoding Rhodococcus L-phenylalanine dehydrogenase - for  
 PT production of recombinant enzyme  
 PT  
 PS Disclosure: fig 5: 30pp: English.  
 XX  
 XX This sequence is used to describe a method which results in the isolation  
 CC of a recombinant L-phenylalanine dehydrogenase gene from Rhodococcus  
 CC The encoding pph nucleic acid is used to produce recombinant Rhodococcus  
 CC L-phenylalanine dehydrogenase, which can be used to produce L-phenylalanine  
 CC from the corresponding alpha-keto acids, e.g. to produce L-phenylalanine  
 CC from phenylpyruvic acid (e.g. for the production of aspartame), to screen  
 CC for phenylketonuria in newborn infants and to monitor the efficacy of  
 CC low-phenylalanine diets.  
 XX  
 SQ Sequence 95 AA:

Query Match 90.9%; Score 30; DB 20; Length 95;  
 Best Local Similarity 83.3%; Pred. No. 4, e+02;  
 Matches 57; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCGTSG 6  
 Db 67 g99ftg 72

RESULT 39  
 ID AAG27159 standard; Protein: 116 AA.  
 AC AAG27159:  
 XX 17-OCT-2000 (first entry)  
 DT 17-OCT-2000 (first entry)  
 XX  
 XX Zea mays protein fragment SEQ ID NO: 31888.  
 DE  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 OS Zea mays subsp. mays.  
 OS  
 XX EP1034405-A2.  
 PN  
 XX 06-SEP-2000.  
 XX  
 XX 25-FEB-2000: 2000EP-0301439.  
 PF  
 XX 25-FEB-1999: 99US-0121825.  
 XX 05-MAR-1999: 99US-0123180.  
 PR 05-MAR-1999: 99US-0123748.  
 PR 23-MAR-1999: 99US-0125780.

PR 25-MAR-1999: 9905-012264.  
PR 25-MAR-1999: 9905-012785.  
PR 03-MAR-1999: 9905-012823.  
PR 06-APR-1999: 9905-0128234.  
PR 08-APR-1999: 9905-0128714.  
PR 16-APR-1999: 9905-0129845.  
PR 19-APR-1999: 9905-0130077.  
PR 21-APR-1999: 9905-0130449.  
PR 23-APR-1999: 9905-0130510.  
PR 23-APR-1999: 9905-0130891.  
PR 30-APR-1999: 9905-0130848.  
PR 30-APR-1999: 9905-0132407.  
PR 04-MAY-1999: 9905-0132484.  
PR 05-MAY-1999: 9905-0132485.  
PR 06-MAY-1999: 9905-0132486.  
PR 06-MAY-1999: 9905-0132487.  
PR 11-MAY-1999: 9905-0132952.  
PR 14-MAY-1999: 9905-0134218.  
PR 14-MAY-1999: 9905-0134219.  
PR 14-MAY-1999: 9905-0134221.  
PR 14-MAY-1999: 9905-0134370.  
PR 18-MAY-1999: 9905-0134768.  
PR 19-MAY-1999: 9905-0134941.  
PR 19-MAY-1999: 9905-0135154.  
PR 21-MAY-1999: 9905-0135154.  
PR 24-MAY-1999: 9905-0136629.  
PR 25-MAY-1999: 9905-0136621.  
PR 27-MAY-1999: 9905-0136392.  
PR 28-MAY-1999: 9905-0136782.  
PR 01-JUN-1999: 9905-0137222.  
PR 04-JUN-1999: 9905-0137225.  
PR 07-JUN-1999: 9905-0137724.  
PR 08-JUN-1999: 9905-0138094.  
PR 10-JUN-1999: 9905-0138540.  
PR 10-JUN-1999: 9905-0138847.  
PR 14-JUN-1999: 9905-0139119.  
PR 16-JUN-1999: 9905-0139432.  
PR 17-JUN-1999: 9905-0139432.  
PR 18-JUN-1999: 9905-0139454.  
PR 18-JUN-1999: 9905-0139455.  
PR 18-JUN-1999: 9905-0139456.  
PR 18-JUN-1999: 9905-0139457.  
PR 18-JUN-1999: 9905-0139458.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139461.  
PR 18-JUN-1999: 9905-0139462.  
PR 18-JUN-1999: 9905-0139463.  
PR 18-JUN-1999: 9905-0139750.  
PR 18-JUN-1999: 9905-0139763.  
PR 21-JUN-1999: 9905-0139817.  
PR 21-JUN-1999: 9905-0139817.  
PR 23-JUN-1999: 9905-0140353.  
PR 23-JUN-1999: 9905-0140354.  
PR 24-JUN-1999: 9905-0140695.  
PR 28-JUN-1999: 9905-0140823.  
PR 29-JUN-1999: 9905-0140991.  
PR 30-JUN-1999: 9905-0141287.  
PR 01-JUL-1999: 9905-0141562.  
PR 01-JUL-1999: 9905-0141562.  
PR 02-JUL-1999: 9905-0142055.  
PR 06-JUL-1999: 9905-0142390.  
PR 08-JUL-1999: 9905-0142803.  
PR 09-JUL-1999: 9905-0142920.  
PR 12-JUL-1999: 9905-0142977.  
PR 14-JUL-1999: 9905-0143632.  
PR 15-JUL-1999: 9905-0144005.  
PR 16-JUL-1999: 9905-0144085.  
  
PR 16-JUL-1999: 9905-0144086.  
PR 19-JUL-1999: 9905-0144335.  
PR 19-JUL-1999: 9905-0144331.  
PR 19-JUL-1999: 9905-0144332.  
PR 19-JUL-1999: 9905-0144333.  
PR 19-JUL-1999: 9905-0144334.  
PR 19-JUL-1999: 9905-0144335.  
PR 20-JUL-1999: 9905-0144332.  
PR 20-JUL-1999: 9905-0144652.  
PR 21-JUL-1999: 9905-0144884.  
PR 21-JUL-1999: 9905-0144884.  
PR 21-JUL-1999: 9905-0145088.  
PR 22-JUL-1999: 9905-0145085.  
PR 22-JUL-1999: 9905-0145087.  
PR 22-JUL-1999: 9905-0145089.  
PR 23-JUL-1999: 9905-0145192.  
PR 23-JUL-1999: 9905-0145115.  
PR 23-JUL-1999: 9905-0145218.  
PR 26-JUL-1999: 9905-0145276.  
PR 27-JUL-1999: 9905-0145913.  
PR 27-JUL-1999: 9905-0145918.  
PR 28-JUL-1999: 9905-0145951.  
PR 02-AUG-1999: 9905-0146386.  
PR 02-AUG-1999: 9905-0146388.  
PR 03-AUG-1999: 9905-0147038.  
PR 04-AUG-1999: 9905-0147204.  
PR 04-AUG-1999: 9905-0147302.  
PR 05-AUG-1999: 9905-0147180.  
PR 06-AUG-1999: 9905-0147260.  
PR 06-AUG-1999: 9905-0147263.  
PR 09-AUG-1999: 9905-0147493.  
PR 09-AUG-1999: 9905-0147955.  
PR 11-AUG-1999: 9905-0148171.  
PR 11-AUG-1999: 9905-0148319.  
PR 12-AUG-1999: 9905-0148342.  
PR 13-AUG-1999: 9905-0148555.  
PR 13-AUG-1999: 9905-0148684.  
PR 15-AUG-1999: 9905-0149175.  
PR 17-AUG-1999: 9905-0149175.  
PR 18-AUG-1999: 9905-0149456.  
PR 20-AUG-1999: 9905-0149722.  
PR 20-AUG-1999: 9905-0149723.  
PR 20-AUG-1999: 9905-0149929.  
PR 23-AUG-1999: 9905-0149902.  
PR 23-AUG-1999: 9905-0150056.  
PR 25-AUG-1999: 9905-0150566.  
PR 26-AUG-1999: 9905-0150884.  
PR 27-AUG-1999: 9905-0151065.  
PR 27-AUG-1999: 9905-0151066.  
PR 27-AUG-1999: 9905-0151080.  
PR 30-AUG-1999: 9905-0151303.  
PR 01-SEP-1999: 9905-0151458.  
PR 01-SEP-1999: 9905-0151458.  
PR 07-SEP-1999: 9905-0151563.  
PR 10-SEP-1999: 9905-0153070.  
PR 13-SEP-1999: 9905-0153758.  
PR 15-SEP-1999: 9905-0154018.  
PR 16-SEP-1999: 9905-0154039.  
PR 20-SEP-1999: 9905-0154773.  
PR 23-SEP-1999: 9905-0155486.  
PR 24-SEP-1999: 9905-0155659.  
PR 28-SEP-1999: 9905-0156458.  
PR 29-SEP-1999: 9905-0156556.  
PR 04-OCT-1999: 9905-0157117.  
PR 05-OCT-1999: 9905-0157283.  
PR 05-OCT-1999: 9905-0157283.  
PR 07-OCT-1999: 9905-0158093.

PR 08-OCT-1999: 9905-018222.  
PR 11-OCT-1999: 9905-018223.  
PR 11-OCT-1999: 9905-018225.  
PR 13-OCT-1999: 9905-0159283.  
PR 13-OCT-1999: 9905-0159285.  
PR 14-OCT-1999: 9905-0159329.  
PR 14-OCT-1999: 9905-0159330.  
PR 14-OCT-1999: 9905-0159331.  
PR 14-OCT-1999: 9905-0159337.  
PR 14-OCT-1999: 9905-0159338.  
PR 18-OCT-1999: 9905-0159584.  
PR 21-OCT-1999: 9905-0160741.  
PR 21-OCT-1999: 9905-0160767.  
PR 21-OCT-1999: 9905-0160768.  
PR 21-OCT-1999: 9905-0160770.  
PR 21-OCT-1999: 9905-0160814.  
PR 21-OCT-1999: 9905-0160815.  
PR 22-OCT-1999: 9905-0160816.  
PR 22-OCT-1999: 9905-0160981.  
PR 22-OCT-1999: 9905-0160989.  
PR 25-OCT-1999: 9905-0161404.  
PR 25-OCT-1999: 9905-0161405.  
PR 25-OCT-1999: 9905-0161406.  
PR 26-OCT-1999: 9905-0161359.  
PR 26-OCT-1999: 9905-0161360.  
PR 26-OCT-1999: 9905-0161361.  
PR 28-OCT-1999: 9905-0161920.  
PR 28-OCT-1999: 9905-0161992.  
PR 28-OCT-1999: 9905-0161993.  
PR 29-OCT-1999: 9905-0162142.

Query Match  
Query Local Similarity 99.98: Score 30. DB 21: Length 116:  
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GGGTGG 6  
DB 37 gggtag 42

RESULT 40

AC06217  
ID AAC06217 standard: Protein: 256 AA.

XX AAC06217:

XX 17-OCT-2000 (first entry)

XX Arabidopsis thaliana protein fragment SEQ ID NO: 2910.

KW Protein identification: signal transduction pathway: metabolic pathway:

KW Hybridisation assay: genetic mapping: gene expression control: promoter:

KW Termination sequence.

OS Arabidopsis thaliana.

XX EPI033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000: 2000EP-0301439.

XX 25-FEB-1999: 9905-0121825.

XX 03-MAR-1999: 9905-0121830.

XX 23-MAR-1999: 9905-0125788.

XX 25-MAR-1999: 9905-0126564.

XX 29-MAR-1999: 9905-0126785.

XX 01-APR-1999: 9905-0127462.

XX 06-APR-1999: 9905-0128234.

XX 08-APR-1999: 9905-0128714.

XX 16-APR-1999: 9905-0129645.

XX 19-APR-1999: 9905-0130077.

PR 21-APR-1999: 9905-0130449.  
PR 23-APR-1999: 9905-0130510.  
PR 23-APR-1999: 9905-0130691.  
PR 23-APR-1999: 9905-0130692.  
PR 30-APR-1999: 9905-0132048.  
PR 30-APR-1999: 9905-0132407.  
PR 04-MAY-1999: 9905-0132484.  
PR 05-MAY-1999: 9905-0132485.  
PR 05-MAY-1999: 9905-0132486.  
PR 06-MAY-1999: 9905-0132487.  
PR 07-MAY-1999: 9905-0132488.  
PR 11-MAY-1999: 9905-0132543.  
PR 14-MAY-1999: 9905-0134218.  
PR 14-MAY-1999: 9905-0134219.  
PR 14-MAY-1999: 9905-0134221.  
PR 14-MAY-1999: 9905-0134370.  
PR 14-MAY-1999: 9905-0134768.  
PR 15-MAY-1999: 9905-0134941.  
PR 15-MAY-1999: 9905-0135343.  
PR 21-MAY-1999: 9905-0135344.  
PR 24-MAY-1999: 9905-0135629.  
PR 25-MAY-1999: 9905-0136021.  
PR 27-MAY-1999: 9905-0136392.  
PR 28-MAY-1999: 9905-0136782.  
PR 01-JUN-1999: 9905-0137222.  
PR 01-JUN-1999: 9905-0137228.  
PR 01-JUN-1999: 9905-0137229.  
PR 07-JUN-1999: 9905-0137722.  
PR 08-JUN-1999: 9905-0138094.  
PR 10-JUN-1999: 9905-0138540.  
PR 10-JUN-1999: 9905-0138847.  
PR 14-JUN-1999: 9905-0139119.  
PR 15-JUN-1999: 9905-0139452.  
PR 15-JUN-1999: 9905-0139453.  
PR 15-JUN-1999: 9905-0139454.  
PR 16-JUN-1999: 9905-0139454.  
PR 18-JUN-1999: 9905-0139455.  
PR 18-JUN-1999: 9905-0139456.  
PR 18-JUN-1999: 9905-0139457.  
PR 18-JUN-1999: 9905-0139458.  
PR 18-JUN-1999: 9905-0139459.  
PR 18-JUN-1999: 9905-0139460.  
PR 18-JUN-1999: 9905-0139461.  
PR 18-JUN-1999: 9905-0139462.  
PR 18-JUN-1999: 9905-0139463.  
PR 18-JUN-1999: 9905-0139750.  
PR 18-JUN-1999: 9905-0139763.  
PR 21-JUN-1999: 9905-0139817.  
PR 22-JUN-1999: 9905-0139899.  
PR 23-JUN-1999: 9905-0140321.  
PR 23-JUN-1999: 9905-0140323.  
PR 24-JUN-1999: 9905-0140695.  
PR 28-JUN-1999: 9905-0140823.  
PR 29-JUN-1999: 9905-0140991.  
PR 30-JUN-1999: 9905-0141287.  
PR 01-JUL-1999: 9905-0141842.  
PR 01-JUL-1999: 9905-0142154.  
PR 06-JUL-1999: 9905-0142925.  
PR 06-JUL-1999: 9905-0142803.  
PR 08-JUL-1999: 9905-0142803.  
PR 09-JUL-1999: 9905-0142920.  
PR 12-JUL-1999: 9905-0142977.  
PR 13-JUL-1999: 9905-0143542.  
PR 14-JUL-1999: 9905-0136324.  
PR 15-JUL-1999: 9905-0144005.  
PR 15-JUL-1999: 9905-0144082.  
PR 15-JUL-1999: 9905-0144321.  
PR 15-JUL-1999: 9905-0144322.  
PR 19-JUL-1999: 9905-0144331.  
PR 19-JUL-1999: 9905-0144332.  
PR 19-JUL-1999: 9905-0144333.  
PR 19-JUL-1999: 9905-0144334.  
PR 19-JUL-1999: 9905-0144335.

PR 20-JUL-1999: 9905-0144352.  
 PR 20-JUL-1999: 9905-0144632.  
 PR 20-JUL-1999: 9905-0144884.  
 PR 20-JUL-1999: 9905-0144944.  
 PR 21-JUL-1999: 9905-0145088.  
 PR 21-JUL-1999: 9905-0145088.  
 PR 22-JUL-1999: 9905-0145087.  
 PR 22-JUL-1999: 9905-0145089.  
 PR 22-JUL-1999: 9905-0145192.  
 PR 23-JUL-1999: 9905-0145218.  
 PR 23-JUL-1999: 9905-0145224.  
 PR 26-JUL-1999: 9905-0145276.  
 PR 27-JUL-1999: 9905-0145918.  
 PR 27-JUL-1999: 9905-0145919.  
 PR 02-AUG-1999: 9905-0145951.  
 PR 02-AUG-1999: 9905-0146388.  
 PR 02-AUG-1999: 9905-0146389.  
 PR 03-AUG-1999: 9905-0147038.  
 PR 04-AUG-1999: 9905-0147204.  
 PR 04-AUG-1999: 9905-0147302.  
 PR 05-AUG-1999: 9905-0147192.  
 PR 06-AUG-1999: 9905-0147260.  
 PR 06-AUG-1999: 9905-0147416.  
 PR 09-AUG-1999: 9905-0147493.  
 PR 09-AUG-1999: 9905-0147935.  
 PR 10-AUG-1999: 9905-0148171.  
 PR 11-AUG-1999: 9905-0148319.  
 PR 12-AUG-1999: 9905-0148241.  
 PR 13-AUG-1999: 9905-0148684.  
 PR 16-AUG-1999: 9905-0149368.  
 PR 17-AUG-1999: 9905-0149175.  
 PR 18-AUG-1999: 9905-0149426.  
 PR 20-AUG-1999: 9905-0149722.  
 PR 20-AUG-1999: 9905-0149923.  
 PR 23-AUG-1999: 9905-0149902.  
 PR 23-AUG-1999: 9905-0149930.  
 PR 25-AUG-1999: 9905-0150566.  
 PR 26-AUG-1999: 9905-0150884.  
 PR 27-AUG-1999: 9905-0151065.  
 PR 27-AUG-1999: 9905-0151066.  
 PR 30-AUG-1999: 9905-0151203.  
 PR 31-AUG-1999: 9905-0151438.  
 PR 01-SEP-1999: 9905-0151930.  
 PR 07-SEP-1999: 9905-0152363.  
 PR 10-SEP-1999: 9905-0153070.  
 PR 13-SEP-1999: 9905-0153758.  
 PR 12-SEP-1999: 9905-0154318.  
 PR 20-SEP-1999: 9905-0154779.  
 PR 22-SEP-1999: 9905-0155139.  
 PR 23-SEP-1999: 9905-0155486.  
 PR 24-SEP-1999: 9905-0155659.  
 PR 28-SEP-1999: 9905-0156458.  
 PR 29-SEP-1999: 9905-0156596.  
 PR 05-OCT-1999: 9905-0157153.  
 PR 06-OCT-1999: 9905-0157865.  
 PR 07-OCT-1999: 9905-0158029.  
 PR 08-OCT-1999: 9905-0158232.  
 PR 12-OCT-1999: 9905-0158369.  
 PR 13-OCT-1999: 9905-0159293.  
 PR 13-OCT-1999: 9905-0159325.  
 PR 14-OCT-1999: 9905-0159329.  
 PR 14-OCT-1999: 9905-0159330.

PR 14-OCT-1999: 9905-0159331.  
 PR 14-OCT-1999: 9905-0159637.  
 PR 18-OCT-1999: 9905-0159684.  
 PR 21-OCT-1999: 9905-0159588.  
 PR 21-OCT-1999: 9905-0160741.  
 PR 21-OCT-1999: 9905-0160767.  
 PR 21-OCT-1999: 9905-0160770.  
 PR 21-OCT-1999: 9905-0160814.  
 PR 21-OCT-1999: 9905-0160815.  
 PR 22-OCT-1999: 9905-0160981.  
 PR 22-OCT-1999: 9905-0160989.  
 PR 25-OCT-1999: 9905-0161405.  
 PR 25-OCT-1999: 9905-0161406.  
 PR 26-OCT-1999: 9905-0161359.  
 PR 26-OCT-1999: 9905-0161360.  
 PR 26-OCT-1999: 9905-0161361.  
 PR 28-OCT-1999: 9905-0161920.  
 PR 28-OCT-1999: 9905-0161923.  
 PR 28-OCT-1999: 9905-0162142.  
 PR 29-OCT-1999: 9905-0162142.

Query Match 90.9% Score 30: DB 21: Length 256.  
 Best Local Similarity 83.3%: Pred.No. 118+03:  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
 Db 231 ggcgsc 236  
 Oy 1 ggcgsc 6

Search completed: February 4, 2002, 08:03:03  
 Job time: 167 sec





GenCore version 4.5  
(c) 1993 - 2000 Com

ein search, using sw model

February 4, 2002, 08:00:16 ; Search time 12.9 Seconds

17.053 Million cell updates/sec

US-09-642-660-10  
33

GGGTSG 6

BLOSUM62

100059 seqs, 36664827 residues

its satisfying chosen parameters:

length: 0

### Minimum Match 02

maximun match 100%  
listing first 100 summaries

SwissProt 39:\*

**Pred. NO.** is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	33	100.0	102	1	YIN2.MEVL	P05410 methanococcus
2	33	100.0	84	1	CIN4.HUMAN	P01427 homo sapiens
3	33	100.0	78	1	MOL1.PROSD	P04433 prunus dulcis
4	33	100.0	563	1	MOL1.PROSD	P04433 prunus dulcis
5	33	100.0	568	1	PLIB.LACIA	P32100 prunus serotina
6	33	100.0	568	1	PLIB.LACIA	P32100 prunus serotina
7	33	100.0	572	1	PLIB.STRPU	P50976 streptococcus
8	33	100.0	572	1	PLIB.LACIA	P24400 lactobacillus
9	33	100.0	577	1	PLIB.STRPU	P44600 streptococcus
10	33	100.0	655	1	GCKR.HUMAN	P01437 homo sapiens
11	33	100.0	665	1	GCKR.HAT	P07017 ratu
12	33	100.0	659	1	KELC3.HUMAN	P12053 homo sapiens
13	33	100.0	677	1	SRB2.DICHD	P16793 human cyclon
14	33	100.0	677	1	SRB2.DICHD	P16793 human cyclon
15	33	100.0	780	1	NH4B.CAEBL	P04407 caecobacteri
16	33	100.0	756	1	KFC3.RAT	O55165 ratu
17	33	100.0	933	1	CLAB.LYCDS	P13542 lycopersicon
18	33	100.0	936	1	CLAB.LYCDS	P13542 lycopersicon
19	30	90.9	92	1	Y155.CAEBL	P34435 cano
20	30	90.9	105	1	NMD1.ECOLI	P04715 escherichia
21	30	90.9	153	1	HNK1.HUMAN	P17727 human
22	30	90.9	258	1	EUTJ.ECOLI	P72277 escherichia
23	30	90.9	278	1	EUTJ.ECOLI	P72277 escherichia
24	30	90.9	129	1	EUTJ.SALTY	P10544 escherichia
25	30	90.9	358	1	ORX3.BHAE	P09416 brachydan
26	30	90.9	362	1	MURG.STRPU	P09416 brachydan
27	30	90.9	354	1	MURG.STRPU	P09416 brachydan
28	30	90.9	359	1	MAE2.RAT	P58443 mus muscu
29	30	90.9	372	1	DNA.BACSE	P58443 mus muscu
30	30	90.9	382	1	LXMA.MESU	P04650 bacillus st
31	30	90.9	395	1	SRCF.YEAST	P33350 saccharomy
32	30	90.9	396	1	DOL.TREPA	O84676 trepennia p
33	30	90.9	407	1	MURG.YCYCLE	P06552 mycobacteri

35	30	90.9	4.0	1	MORG_MCTD
36	30	90.9	4.0	1	MORG_MCTD
37	30	90.9	4.0	1	MORG_MCTD
38	30	90.9	4.0	1	MORG_MCTD
39	30	90.9	4.0	1	MORG_MCTD
40	30	90.9	4.0	1	MORG_MCTD
41	30	90.9	4.0	1	MORG_MCTD
42	30	90.9	4.0	1	MORG_MCTD
43	30	90.9	4.0	1	MORG_MCTD
44	30	90.9	4.0	1	MORG_MCTD
45	30	90.9	4.0	1	MORG_MCTD
46	30	90.9	4.0	1	MORG_MCTD
47	30	90.9	4.0	1	MORG_MCTD
48	30	90.9	4.0	1	MORG_MCTD
49	30	90.9	4.0	1	MORG_MCTD
50	30	90.9	4.0	1	MORG_MCTD
51	30	90.9	4.0	1	MORG_MCTD
52	30	90.9	4.0	1	MORG_MCTD
53	30	90.9	4.0	1	MORG_MCTD
54	30	90.9	4.0	1	MORG_MCTD
55	30	90.9	4.0	1	MORG_MCTD
56	30	90.9	4.0	1	MORG_MCTD
57	30	90.9	4.0	1	MORG_MCTD
58	30	90.9	4.0	1	MORG_MCTD
59	30	90.9	4.0	1	MORG_MCTD
60	30	90.9	4.0	1	MORG_MCTD
61	30	90.9	4.0	1	MORG_MCTD
62	30	90.9	4.0	1	MORG_MCTD
63	30	90.9	4.0	1	MORG_MCTD
64	30	90.9	4.0	1	MORG_MCTD
65	30	90.9	4.0	1	MORG_MCTD
66	30	90.9	4.0	1	MORG_MCTD
67	30	90.9	4.0	1	MORG_MCTD
68	30	90.9	4.0	1	MORG_MCTD
69	30	90.9	4.0	1	MORG_MCTD
70	30	90.9	4.0	1	MORG_MCTD
71	30	90.9	4.0	1	MORG_MCTD
72	30	90.9	4.0	1	MORG_MCTD
73	30	90.9	4.0	1	MORG_MCTD
74	30	90.9	4.0	1	MORG_MCTD
75	30	90.9	4.0	1	MORG_MCTD
76	30	90.9	4.0	1	MORG_MCTD
77	30	90.9	4.0	1	MORG_MCTD
78	30	90.9	4.0	1	MORG_MCTD
79	30	90.9	4.0	1	MORG_MCTD
80	30	90.9	4.0	1	MORG_MCTD
81	30	90.9	4.0	1	MORG_MCTD
82	30	90.9	4.0	1	MORG_MCTD
83	30	90.9	4.0	1	MORG_MCTD
84	30	90.9	4.0	1	MORG_MCTD
85	30	90.9	4.0	1	MORG_MCTD
86	30	90.9	4.0	1	MORG_MCTD
87	30	90.9	4.0	1	MORG_MCTD
88	30	90.9	4.0	1	MORG_MCTD
89	30	90.9	4.0	1	MORG_MCTD
90	30	90.9	4.0	1	MORG_MCTD
91	30	90.9	4.0	1	MORG_MCTD
92	30	90.9	4.0	1	MORG_MCTD
93	30	90.9	4.0	1	MORG_MCTD
94	30	90.9	4.0	1	MORG_MCTD
95	30	90.9	4.0	1	MORG_MCTD
96	30	90.9	4.0	1	MORG_MCTD
97	30	90.9	4.0	1	MORG_MCTD
98	30	90.9	4.0	1	MORG_MCTD
99	30	90.9	4.0	1	MORG_MCTD
100	30	90.9	4.0	1	MORG_MCTD

```

NT2.METTL.          STANDARD:      PRT: 102 AA.
AD P05410
DT 01-NOV-1988 (rel. 09, Created)
DT 01-NOV-1988 (rel. 09, Last sequence update)
DT 01-NOV-1990 (rel. 16, Last annotation update)
DE HYPOTHETICAL PROTEIN IN NIFH2.3 REGION (FRAGMENT).
OS Methanococcus thermophilus.
CC Archaea: Euryarchaeota: Methanococcales: Methanococcaceae:
OC Methanococcus thermophilus.
CX NCBI_TaxID=2186;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=88259240; PubMed=2838639;
RA Soullard N., Magot M., Posset O., Sibold L.;
RT Nucleotide sequence of regions homologous to nifH (nitrogenase Fe
RT protein) from the nitrogen-fixing archaebacteria Methanococcus
RT thermophilus and Methanobacterium Yarrowii: evolutionary
RT implications.
RL J. Mol. Evol. 27:65-76(1988).
CC
CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use for non-profit institutions as long as its content is in no way
CC used for profit making, and no fee is charged by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: X07500; CNA30382.1;
DR PIR: S00739; S00739.
DR InterPro: IPR000205; NMD-binding.
DR KEGG: 102; Protein:102
DR NON-REF: 102
SQ SEQUENCE 102 AA: 11362 MW: 140266961975FE2 CRC64:
Query Match 100.0%; Score 33; DB 1; Length 102;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTSG 6
DB 7 GCGTSG 12
ID COX4_HUMAN STANDARD: PRT: 284 AA.
AC 014627;
DT 15-JUL-1998 (rel. 36, Created)
DT 15-JUL-1998 (rel. 36, Last sequence update)
DT 30-MAY-2000 (rel. 39, Last annotation update)
DE CYTOCHROME C OXIDASE SUBUNIT 4 (CAUDAL-TYPE HOMODIOL PROTEIN 4).
OS Homo sapiens (Human).
OC Eukaryota: Metazoa: Chordata: Canitata: Vertebrata: Euteleostomi:
OC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=152222;
RA Shen E., States D.J., Mazarella R.;
RT Submitted (MAY-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Chelaron P., Colletaux L., Brelaut S., Losal A.M., Moraine C.;
RT Genomic structure of the human COX4 gene: a potential candidate for
RT the PC syndrome.
RL Submitted (OCT-1997) to the EMBL/Genbank/DBJ databases.
CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use for non-profit institutions as long as its content is in no way
CC used for profit making, and no fee is charged by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----

```

```

CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use for non-profit institutions as long as its content is in no way
CC used for profit making, and no fee is charged by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: AF003530; AAB6319.1;
DR EMBL: AF029875; AAD01894.1; UNIND.
DR EMBL: AF029875; AAD01894.1; JOINED.
DR HSPR: P02835; IFTZ.
DR MKM: 300025;
DR InterPro: IPR001827; Antennapedia.
DR InterPro: IPR000047; HTH_repressor.
DR InterPro: IPR001356; Homocobox.
DR Pfam: PF00046; Homocobox.1.
DR PRINTS: PR00025; ANTENNAPEDIA.
DR PRINTS: PR00031; HTHREPRESSR.
DR SMART: SM00389; HOX.1.
DR PROSITE: PS00027; HOMODIOL_1; 1.
DR PROSITE: PS50071; HOMODIOL_2; 1.
DR Homocobox: DNA-binding; developmental protein: Nuclear protein:
DR Transcription regulation.
DR KEGG: 284; Protein:284
SQ SEQUENCE 284 AA: 30480 MW: 960385095760607 CRC64:
Query Match 100.0%; Score 33; DB 1; Length 284;
Best Local Similarity 100.0%; Pred. No. 57;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCGTSG 6
DB 120 GCGTSG 125
ID HMD1_PRUDU STANDARD: PRT: 559 AA.
AC 024243;
DT 15-DEC-1998 (rel. 37, Created)
DT 15-DEC-1998 (rel. 37, Last sequence update)
DT 15-DEC-1998 (rel. 37, Last annotation update)
DE (R)-MANDELONITRILE LIASE ISOFORM 1 PRECURSOR (EC 4.1.2.10)
DE HYDROXYNITRILE LIASE 1 ((R)-OXALNITRILE 1).
OS Prunus dulcis (Almond) (Prunus amygdalus).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:
OX NCBI_TaxID=3755;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=152222;
RA Shen E., States D.J., Mazarella R.;
RT Submitted (SEP-1996) to the EMBL/Genbank/DBJ databases.
RL FUNCTION: INVOLVED IN CYANOGENESIS, THE RELEASE OF HCN FROM INJURED
CC TISSUES. CATALYSES THE STEREOSPECIFIC ADDITION OF HCN TO A VARIETY
CC OF ALDEHYDES IN VITRO. IS A MAJOR SPECIFIC CONSTITUENT, AND COULD HAVE
CC THE ADDITIONAL ROLE OF A STORAGE FORM FOR REDUCED NITROGEN (BY
CC -1- CATALYTIC ACTIVITY: MANDELONITRILE - CYANIDE + BENZALDEHYDE.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.
CC
CC This SMS-S-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use for non-profit institutions as long as its content is in no way
CC used for profit making, and no fee is charged by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----

```

CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).

CC EMBL: U082116; CAA59386.1;  
 CC Protein: P000732; GMC\_OXRED.1; Oxidized.  
 CC DR PROSITE: PS00624; GMC\_OXRED.2; 1.  
 CC DR Lysase: Glycoprotein; FAD; Flavoprotein; Signal; Multigene family.  
 CC SIGNAL 1 27  
 CC FT CHAIN 28 559 (R)-MANDELONITRILE LYASE ISOFORM 1.  
 CC FT PROSITE: PS00732; GMC\_OXRED.1;  
 CC FT CARBOHYD 30 36 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 44 44 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 145 145 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 162 162 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 178 178 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 218 218 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 232 232 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 252 252 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 309 309 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 380 380 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 402 402 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 420 420 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 467 467 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT SEQUENCE 559 AA; 61099 MW; 58C2959BEF28A54 CRC64;

Query Match 100.0%; Score 33; DB 1; Length 559;  
 Best Local Similarity 100.0%; Pred. No. 1;le02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGSTSG 6  
 DB 60 GGSTSG 65

RESULT 4  
 MDL\_PROUSE STANDARD; PRT; 563 AA.  
 ID MDL\_PROUSE  
 AC P52706; 1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE (R)-MANDELONITRILE LYASE ISOFORM 1 PRECURSOR (EC 4.1.2.10)  
 DE (HYDROXYNITRILE LYASE 1) ((R)-OXYNITRILE 1).  
 GN MDL.  
 OS Prunus serotina (Black cherry).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Angiosperms; Eudicotyledons; core eudicots; Rosidae;  
 OC Malvaceae; Malvaceae; Prunus.  
 OC NCBI\_TaxID:23207;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RA TISSUE-Seeds.  
 RC Cheng I.-P., Poulton J.E.;  
 RT Cloning of cDNA of Prunus serotina (R-+) mandelonitrile lyase and  
 RT identification of a homologous gene encoding site-7.  
 RT Plant Cell Physiol. 34:1139-1143(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Hu Z., Poulton J.E.;  
 RT Submitted (DEC-1996) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP TISSUE-Seeds.  
 RC Cheng I.-P., Poulton J.E.;  
 RA \*Immunocytochemical localization of mandelonitrile lyase in mature  
 RA black cherry (Prunus serotina Ehrh.) seeds \*;  
 RT Plant Physiol. 96:1329-1337(1991).  
 CC -1- FUNCTION: INVOLVED IN CYANOGENS, THE RELEASE OF HCN FROM INJURED  
 CC TISSUES. CATALYZES THE STEREOSPECIFIC ADDITION OF HCN TO A VARIETY

CC OF ALDEHYDES IN VITRO. IS A MAJOR SEED CONSTITUENT, AND COULD HAVE  
 CC THE ADDITIONAL ROLE OF A STORAGE FORM FOR REDUCED NITROGEN.  
 CC -1- CATALYTIC ACTIVITY: MANDELONITRILE - CYANIDE + BENZALDEHYDE.  
 CC -1- COFACTOR: FAD.  
 CC -1- SUBCELLULAR LOCATION: CYTOSOL.  
 CC -1- SUBCELLULAR LOCATION: PRIMARY FOUND WITHIN PROTEIN BODIES OF THE  
 CC PROCAMBIIUM.  
 CC -1- TISSUE SPECIFICITY: SEEDS.  
 CC -1- DEVELOPMENTAL STAGE: ABSENT FROM MATURING BLACK CHERRY FRUITS  
 CC UNTIL 6 WEEKS AFTER FLOWERING. THEN, CONCOMITANT WITH COTYLEDON  
 CC DEVELOPMENT, THE LEVEL OF ENZYME INCREASES WITH SPECIFICITY FOR  
 CC PROCAMBIIUM.  
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.  
 CC -1- THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on  
 CC use of this information, and it is made available under the conditions  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).

CC EMBL: X72617; CAA51194.1;  
 CC Protein: U78814; AAB58536.1;  
 CC InterPro: IPR000172; GMC\_OXRED.  
 CC DR PROSITE: PS00624; GMC\_OXRED.2; 1.  
 CC DR Lysase: Glycoprotein; FAD; Flavoprotein; Signal; Multigene family.  
 CC SIGNAL 1 27  
 CC FT CHAIN 28 563 (H)-MANDELONITRILE LYASE ISOFORM 1.  
 CC FT PROSITE: PS00732; GMC\_OXRED.1;  
 CC FT CARBOHYD 35 82 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 145 145 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 162 162 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT CARBOHYD 379 379 N-LINKED (GLCNAC... ) (POTENTIAL).  
 CC FT SEQUENCE 563 AA; 61200 MW; 152367E736A5F5PD CRC64;

Query Match 100.0%; Score 33; DB 1; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 1;le02; Mismatches 0; Indels 0; Gaps 0;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGSTSG 6  
 DB 60 GGSTSG 65

RESULT 5  
 MDL\_LACTA STANDARD; PRT; 566 AA.  
 ID MDL\_LACTA  
 AC P23531;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 20, Last annotation update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE PTS SYSTEM, LACTOSE-SPECIFIC IIIB COMPONENT (ELIHC-LAC) (LACTOSE-  
 DE PERMEASE IIIB COMPONENT) (PHOSPHOTRANSFERASE ENZYME II, BC COMPONENT\*)  
 DE LACE 2.7.1.69) (ELI-LAC).  
 GN LACTA  
 OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).  
 OC Bacteria; Firmicutes; Bacilli/Clostridium group; Streptococcaceae;  
 OC Lactococcus.  
 OC NCBI\_TaxID:1360;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-MG1820.  
 RC MEDLINE-91093107; Pubmed-2125052;  
 RA de Vos M.M., Boerlijter I.J., van Nooyen R.J., Reiche B.,  
 RA Hengstenberg W.;  
 RT Characterization of the lactose-specific enzymes of the  
 RT phosphotransferase system in Lactococcus lactis-7;





ID GCRK\_HUMAN STANDARD: PRT: 625 AA.  
 AC 014397: 099522: 35: Created  
 DT 01-OCT-1994 (rel. 30, Last sequence update)  
 DT 20-MAY-2000 (rel. 39, Last annotation update)  
 DE GLUCOKINASE REGULATORY PROTEIN (GLUCOKINASE REGULATOR).  
 GN GCRK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCHL\_TaxId-9606;  
 RX MEDLINE-96014291: PubMed-8589523;  
 RA Warner J.P., Leek J.P., Intody S., Markham A.F., Bonthron D.T.;  
 RT Cloning and sequencing of rat liver cDNAs encoding the regulatory  
 protein of glucokinase.  
 RL Mamm. Genome 5:532-536(1995).  
 RP PARTIAL SEQUENCE FROM N.A.; REVISIONS, AND VARIANT LEU-446.  
 RX MEDLINE-98234554: PubMed-9570959;  
 RA Hayward B.E., Dunlop N., Intody S., Leek J.P., Markham A.F.,  
 RT Cloning and sequencing of the human glucokinase regulator gene GCRK.\*  
 RL Mamm. Genome 10:1437-1439(1999).  
 CC -1- FUNCTION: INHIBITS GLUCOKINASE BY FORMING AN INACTIVE COMPLEX WITH  
 THIS ENZYME.  
 CC -1- TISSUE SPECIFICITY: FOUND IN LIVER AND PANCREAS. NOT DETECTED IN  
 MUSCLE, BRAIN, HEART, THYMUS, INTESTINE, UTERUS, ADIPOSE TISSUE,  
 KIDNEY, ADRENAL, LUNG OR SPLEEN.  
 CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GCRK SUSPAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.lsb-sib.ch/announcement/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: Z48475; CABB8367.1; -  
 DR EMBL: Y09593; CA70779.2; -  
 DR EMBL: Y09592; CABB1828.1; -  
 DR MIM: 600842; -  
 DR InterPro: IP001741; GCRK.  
 DR Pfam: PF01380; SIS.  
 DR Pfam: PF01380; SIS.  
 DR ProDom: PD025295; GCRK; 1.  
 DR ProSITE: PS01272; GCRK; 1.  
 KW Polymorphism.  
 FT VARIANT 446 446 P -> L.  
 FT SEQUENCE 625 AA: 68685 MW: DE750462AC603C80 CRC64;  
 Query Match 100.0%; Score 33; DB 1; Length 625;  
 Best Local Similarity 100.0%; Pct. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTSG 6  
 DB 106 GGGTSG 111  
 RESULT 10  
 GCRK\_RAT  
 ID GCRK\_RAT STANDARD: PRT: 626 AA.  
 AC 007071:  
 DT 01-OCT-1994 (rel. 30, Created)  
 DT 01-SEP-1998 (rel. 31, Last sequence update)  
 DT 20-MAY-2000 (rel. 39, Last annotation update)  
 DE GLUCOKINASE REGULATORY PROTEIN (GLUCOKINASE REGULATOR).  
 GN GCRK.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCHL\_TaxId-9606;  
 RX MEDLINE-94156054: PubMed-8112473;  
 RA Dethleux M., Vandekerckhove J., Van Schaefingen E.;  
 RT Cloning and sequencing of rat liver cDNAs encoding the regulatory  
 protein of glucokinase.  
 RL Mamm. Genome 5:532-536(1995).  
 RP REVISIONS.  
 RX MEDLINE-94156054: PubMed-8112473;  
 RA Dethleux M., Vandekerckhove J., Van Schaefingen E.;  
 RT Cloning and sequencing of rat liver cDNAs encoding the regulatory  
 protein of glucokinase.  
 RL Mamm. Genome 5:532-536(1995).  
 CC -1- THIS ENZYME.  
 CC -1- FUNCTION: INHIBITS GLUCOKINASE BY FORMING AN INACTIVE COMPLEX WITH  
 THIS ENZYME.  
 CC -1- TISSUE SPECIFICITY: FOUND IN LIVER. NOT DETECTED IN MUSCLE, BRAIN,  
 HEART, TESTIS, INTESTINE OR SPLEEN.  
 CC -1- SIMILARITY: BELONGS TO THE SIS FAMILY. GCRK SUSPAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation  
 at the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.lsb-sib.ch/announcement/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: X64497; CAA18511;  
 DR InterPro: IP001741; GCRK.  
 DR Pfam: PF01380; SIS.  
 DR ProDom: PD025295; GCRK; 1.  
 DR ProSITE: PS01272; GCRK; 1.  
 FT INIT MET 0  
 FT SEQUENCE 626 AA: 68786 MW: 60M6A57F9ACDIEF3 CRC64;  
 Query Match 100.0%; Score 33; DB 1; Length 626;  
 Best Local Similarity 100.0%; Pct. No. 1.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTSG 6  
 DB 105 GGGTSG 110  
 RESULT 11  
 K2C3\_HUMAN  
 ID K2C3\_HUMAN STANDARD: PRT: 629 AA.  
 AC 014397: 099522: 35: Created  
 DT 01-OCT-1994 (rel. 30, Last sequence update)  
 DT 20-MAY-2000 (rel. 39, Last annotation update)  
 DE KERATIN, TYPE II CYTOSKELETAL 3 (CYTOKERATIN 3) (K3) (CK3) (65 KDA  
 CYTOKERATIN).  
 GN KRT3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCHL\_TaxId-9606;  
 RX MEDLINE-87254239: PubMed-2439698;  
 RA Klinge E.M., Sylvestre Y.R., Freedberg I.M., Blumenberg M.;  
 RT Evolution of keratin genes: different protein domains evolve by  
 different pathways.\*

RL J. Mol. Evol. 24:319-329(1987).  
 RP VARIANT MCD IYS-509  
 RA MEDLINE-97315825; PubMed-9171831;  
 RA Irvine A.D., Cordeiro L.D., Swenson O., Swenson B., Moore J.E.,  
 RA Frazer D.G., Smith F.J.D., Knowlton R.G., Christophers E.,  
 RA Rochels R., Uitto J., McLean W.H.I.;  
 RA Mutations in corneal-specific keratin K3 or K12 genes cause  
 RA Meesmans in corneal dystrophy.\*;  
 RA Med. Genet. 116:187-190(1997).  
 CC -1- SHUNTIN-1 IS ASSOCIATED WITH KERATIN 12.  
 CC KERATIN 3 ASSOCIATES WITH KERATIN 12.  
 CC -1- TISSUE SPECIFICITY: CORNEA-SPECIFIC.  
 CC -1- DISEASE: DEFECTS IN KRT3 ARE A CAUSE OF MEESSMAN CORNEAL DYSTROPHY  
 CC (MCD), AN AUTOSOMAL DOMINANT DISEASE THAT CAUSES FRAGILITY OF THE  
 CC ANTERIOR CORNEAL EPITHELIUM. PATIENTS ARE USUALLY ASYMPTOMATIC  
 CC UNTIL ADULTHOOD WHEN RUPTURE OF THE CORNEAL MICROCYSTS MAY CAUSE  
 CC VISION LOSS. PROGRESSIVE CLINICAL SYMPTOMS SUCH AS PHOTOOPHOBIA, CONTACT  
 CC LENS INTOLERANCE AND RECURRENT CORNEAL EPITHELIAL ACUTITY.  
 CC RARELY, SUBEPITHELIAL SCARRING CAUSES IRREGULAR CORNEAL  
 CC ASTIGMATISM AND PERMANENT VISUAL IMPAIRMENT. HISTOLOGICAL  
 CC EXAMINATION SHOWS A DISORGANIZED AND THICKENED EPITHELIUM WITH  
 CC WIDESPREAD CYTOPLASMIC VACUOLATION AND NUMEROUS SMALL, ROUND,  
 CC DEBRIS-LADEN INTRAEPITHELIAL CYSTS.  
 CC -1- MICELLANEOUS: THERE ARE TWO TYPES OF CYTOSKELETAL AND  
 CC MICELLANEOUS: KERATIN 3 AND KERATIN 12 (K3 AND K12) AND II  
 CC (NEUTRAL TO BASIC; 55-70 KDa) AND KERATIN 12 (K12) AND II  
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X05418; CA28899.1; -; ALT-SEQ.  
 DR EMBL: X05419; CA28899.1; -; ALT-SEQ.  
 DR EMBL: X05420; CA28899.1; -; ALT-SEQ.  
 DR EMBL: X05421; CA28899.1; -; ALT-SEQ.  
 DR EMBL: X05422; CA28899.1; -; ALT-SEQ.  
 DR PIR: A29666; A29666.  
 DR MIM: 148043; -;  
 DR MIM: 122100; -;  
 DR InterPro: IPR001664; IF;  
 DR InterPro: IPR001664; Kcatalin-II.  
 DR Pfam: PF00038; filament\_1.  
 DR PRINTS: PR001276; TYPEKERATIN.  
 DR PROSITE: PS00026; IF; 1.  
 DR Intermediate filament; Coiled coil; Keratin; Phosphorylation;  
 DR Disease mutation.  
 FT 1 197 HEAD.  
 FT 2 500  
 FT 3 500  
 FT 4 500  
 FT 5 500  
 FT 6 500  
 FT 7 500  
 FT 8 500  
 FT 9 500  
 FT 10 500  
 FT 11 500  
 FT 12 500  
 FT 13 500  
 FT 14 500  
 FT 15 500  
 FT 16 500  
 FT 17 500  
 FT 18 500  
 FT 19 500  
 FT 20 500  
 FT 21 500  
 FT 22 500  
 FT 23 500  
 FT 24 500  
 FT 25 500  
 FT 26 500  
 FT 27 500  
 FT 28 500  
 FT 29 500  
 FT 30 500  
 FT 31 500  
 FT 32 500  
 FT 33 500  
 FT 34 500  
 FT 35 500  
 FT 36 500  
 FT 37 500  
 FT 38 500  
 FT 39 500  
 FT 40 500  
 FT 41 500  
 FT 42 500  
 FT 43 500  
 FT 44 500  
 FT 45 500  
 FT 46 500  
 FT 47 500  
 FT 48 500  
 FT 49 500  
 FT 50 500  
 FT 51 500  
 FT 52 500  
 FT 53 500  
 FT 54 500  
 FT 55 500  
 FT 56 500  
 FT 57 500  
 FT 58 500  
 FT 59 500  
 FT 60 500  
 FT 61 500  
 FT 62 500  
 FT 63 500  
 FT 64 500  
 FT 65 500  
 FT 66 500  
 FT 67 500  
 FT 68 500  
 FT 69 500  
 FT 70 500  
 FT 71 500  
 FT 72 500  
 FT 73 500  
 FT 74 500  
 FT 75 500  
 FT 76 500  
 FT 77 500  
 FT 78 500  
 FT 79 500  
 FT 80 500  
 FT 81 500  
 FT 82 500  
 FT 83 500  
 FT 84 500  
 FT 85 500  
 FT 86 500  
 FT 87 500  
 FT 88 500  
 FT 89 500  
 FT 90 500  
 FT 91 500  
 FT 92 500  
 FT 93 500  
 FT 94 500  
 FT 95 500  
 FT 96 500  
 FT 97 500  
 FT 98 500  
 FT 99 500  
 FT 100 500  
 FT 101 500  
 FT 102 500  
 FT 103 500  
 FT 104 500  
 FT 105 500  
 FT 106 500  
 FT 107 500  
 FT 108 500  
 FT 109 500  
 FT 110 500  
 FT 111 500  
 FT 112 500  
 FT 113 500  
 FT 114 500  
 FT 115 500  
 FT 116 500  
 FT 117 500  
 FT 118 500  
 FT 119 500  
 FT 120 500  
 FT 121 500  
 FT 122 500  
 FT 123 500  
 FT 124 500  
 FT 125 500  
 FT 126 500  
 FT 127 500  
 FT 128 500  
 FT 129 500  
 FT 130 500  
 FT 131 500  
 FT 132 500  
 FT 133 500  
 FT 134 500  
 FT 135 500  
 FT 136 500  
 FT 137 500  
 FT 138 500  
 FT 139 500  
 FT 140 500  
 FT 141 500  
 FT 142 500  
 FT 143 500  
 FT 144 500  
 FT 145 500  
 FT 146 500  
 FT 147 500  
 FT 148 500  
 FT 149 500  
 FT 150 500  
 FT 151 500  
 FT 152 500  
 FT 153 500  
 FT 154 500  
 FT 155 500  
 FT 156 500  
 FT 157 500  
 FT 158 500  
 FT 159 500  
 FT 160 500  
 FT 161 500  
 FT 162 500  
 FT 163 500  
 FT 164 500  
 FT 165 500  
 FT 166 500  
 FT 167 500  
 FT 168 500  
 FT 169 500  
 FT 170 500  
 FT 171 500  
 FT 172 500  
 FT 173 500  
 FT 174 500  
 FT 175 500  
 FT 176 500  
 FT 177 500  
 FT 178 500  
 FT 179 500  
 FT 180 500  
 FT 181 500  
 FT 182 500  
 FT 183 500  
 FT 184 500  
 FT 185 500  
 FT 186 500  
 FT 187 500  
 FT 188 500  
 FT 189 500  
 FT 190 500  
 FT 191 500  
 FT 192 500  
 FT 193 500  
 FT 194 500  
 FT 195 500  
 FT 196 500  
 FT 197 500  
 FT 198 500  
 FT 199 500  
 FT 200 500  
 FT 201 500  
 FT 202 500  
 FT 203 500  
 FT 204 500  
 FT 205 500  
 FT 206 500  
 FT 207 500  
 FT 208 500  
 FT 209 500  
 FT 210 500  
 FT 211 500  
 FT 212 500  
 FT 213 500  
 FT 214 500  
 FT 215 500  
 FT 216 500  
 FT 217 500  
 FT 218 500  
 FT 219 500  
 FT 220 500  
 FT 221 500  
 FT 222 500  
 FT 223 500  
 FT 224 500  
 FT 225 500  
 FT 226 500  
 FT 227 500  
 FT 228 500  
 FT 229 500  
 FT 230 500  
 FT 231 500  
 FT 232 500  
 FT 233 500  
 FT 234 500  
 FT 235 500  
 FT 236 500  
 FT 237 500  
 FT 238 500  
 FT 239 500  
 FT 240 500  
 FT 241 500  
 FT 242 500  
 FT 243 500  
 FT 244 500  
 FT 245 500  
 FT 246 500  
 FT 247 500  
 FT 248 500  
 FT 249 500  
 FT 250 500  
 FT 251 500  
 FT 252 500  
 FT 253 500  
 FT 254 500  
 FT 255 500  
 FT 256 500  
 FT 257 500  
 FT 258 500  
 FT 259 500  
 FT 260 500  
 FT 261 500  
 FT 262 500  
 FT 263 500  
 FT 264 500  
 FT 265 500  
 FT 266 500  
 FT 267 500  
 FT 268 500  
 FT 269 500  
 FT 270 500  
 FT 271 500  
 FT 272 500  
 FT 273 500  
 FT 274 500  
 FT 275 500  
 FT 276 500  
 FT 277 500  
 FT 278 500  
 FT 279 500  
 FT 280 500  
 FT 281 500  
 FT 282 500  
 FT 283 500  
 FT 284 500  
 FT 285 500  
 FT 286 500  
 FT 287 500  
 FT 288 500  
 FT 289 500  
 FT 290 500  
 FT 291 500  
 FT 292 500  
 FT 293 500  
 FT 294 500  
 FT 295 500  
 FT 296 500  
 FT 297 500  
 FT 298 500  
 FT 299 500  
 FT 300 500  
 FT 301 500  
 FT 302 500  
 FT 303 500  
 FT 304 500  
 FT 305 500  
 FT 306 500  
 FT 307 500  
 FT 308 500  
 FT 309 500  
 FT 310 500  
 FT 311 500  
 FT 312 500  
 FT 313 500  
 FT 314 500  
 FT 315 500  
 FT 316 500  
 FT 317 500  
 FT 318 500  
 FT 319 500  
 FT 320 500  
 FT 321 500  
 FT 322 500  
 FT 323 500  
 FT 324 500  
 FT 325 500  
 FT 326 500  
 FT 327 500  
 FT 328 500  
 FT 329 500  
 FT 330 500  
 FT 331 500  
 FT 332 500  
 FT 333 500  
 FT 334 500  
 FT 335 500  
 FT 336 500  
 FT 337 500  
 FT 338 500  
 FT 339 500  
 FT 340 500  
 FT 341 500  
 FT 342 500  
 FT 343 500  
 FT 344 500  
 FT 345 500  
 FT 346 500  
 FT 347 500  
 FT 348 500  
 FT 349 500  
 FT 350 500  
 FT 351 500  
 FT 352 500  
 FT 353 500  
 FT 354 500  
 FT 355 500  
 FT 356 500  
 FT 357 500  
 FT 358 500  
 FT 359 500  
 FT 360 500  
 FT 361 500  
 FT 362 500  
 FT 363 500  
 FT 364 500  
 FT 365 500  
 FT 366 500  
 FT 367 500  
 FT 368 500  
 FT 369 500  
 FT 370 500  
 FT 371 500  
 FT 372 500  
 FT 373 500  
 FT 374 500  
 FT 375 500  
 FT 376 500  
 FT 377 500  
 FT 378 500  
 FT 379 500  
 FT 380 500  
 FT 381 500  
 FT 382 500  
 FT 383 500  
 FT 384 500  
 FT 385 500  
 FT 386 500  
 FT 387 500  
 FT 388 500  
 FT 389 500  
 FT 390 500  
 FT 391 500  
 FT 392 500  
 FT 393 500  
 FT 394 500  
 FT 395 500  
 FT 396 500  
 FT 397 500  
 FT 398 500  
 FT 399 500  
 FT 400 500  
 FT 401 500  
 FT 402 500  
 FT 403 500  
 FT 404 500  
 FT 405 500  
 FT 406 500  
 FT 407 500  
 FT 408 500  
 FT 409 500  
 FT 410 500  
 FT 411 500  
 FT 412 500  
 FT 413 500  
 FT 414 500  
 FT 415 500  
 FT 416 500  
 FT 417 500  
 FT 418 500  
 FT 419 500  
 FT 420 500  
 FT 421 500  
 FT 422 500  
 FT 423 500  
 FT 424 500  
 FT 425 500  
 FT 426 500  
 FT 427 500  
 FT 428 500  
 FT 429 500  
 FT 430 500  
 FT 431 500  
 FT 432 500  
 FT 433 500  
 FT 434 500  
 FT 435 500  
 FT 436 500  
 FT 437 500  
 FT 438 500  
 FT 439 500  
 FT 440 500  
 FT 441 500  
 FT 442 500  
 FT 443 500  
 FT 444 500  
 FT 445 500  
 FT 446 500  
 FT 447 500  
 FT 448 500  
 FT 449 500  
 FT 450 500  
 FT 451 500  
 FT 452 500  
 FT 453 500  
 FT 454 500  
 FT 455 500  
 FT 456 500  
 FT 457 500  
 FT 458 500  
 FT 459 500  
 FT 460 500  
 FT 461 500  
 FT 462 500  
 FT 463 500  
 FT 464 500  
 FT 465 500  
 FT 466 500  
 FT 467 500  
 FT 468 500  
 FT 469 500  
 FT 470 500  
 FT 471 500  
 FT 472 500  
 FT 473 500  
 FT 474 500  
 FT 475 500  
 FT 476 500  
 FT 477 500  
 FT 478 500  
 FT 479 500  
 FT 480 500  
 FT 481 500  
 FT 482 500  
 FT 483 500  
 FT 484 500  
 FT 485 500  
 FT 486 500  
 FT 487 500  
 FT 488 500  
 FT 489 500  
 FT 490 500  
 FT 491 500  
 FT 492 500  
 FT 493 500  
 FT 494 500  
 FT 495 500  
 FT 496 500  
 FT 497 500  
 FT 498 500  
 FT 499 500  
 FT 500 500  
 FT 501 500  
 FT 502 500  
 FT 503 500  
 FT 504 500  
 FT 505 500  
 FT 506 500  
 FT 507 500  
 FT 508 500  
 FT 509 500  
 FT 510 500  
 FT 511 500  
 FT 512 500  
 FT 513 500  
 FT 514 500  
 FT 515 500  
 FT 516 500  
 FT 517 500  
 FT 518 500  
 FT 519 500  
 FT 520 500  
 FT 521 500  
 FT 522 500  
 FT 523 500  
 FT 524 500  
 FT 525 500  
 FT 526 500  
 FT 527 500  
 FT 528 500  
 FT 529 500  
 FT 530 500  
 FT 531 500  
 FT 532 500  
 FT 533 500  
 FT 534 500  
 FT 535 500  
 FT 536 500  
 FT 537 500  
 FT 538 500  
 FT 539 500  
 FT 540 500  
 FT 541 500  
 FT 542 500  
 FT 543 500  
 FT 544 500  
 FT 545 500  
 FT 546 500  
 FT 547 500  
 FT 548 500  
 FT 549 500  
 FT 550 500  
 FT 551 500  
 FT 552 500  
 FT 553 500  
 FT 554 500  
 FT 555 500  
 FT 556 500  
 FT 557 500  
 FT 558 500  
 FT 559 500  
 FT 560 500  
 FT 561 500  
 FT 562 500  
 FT 563 500  
 FT 564 500  
 FT 565 500  
 FT 566 500  
 FT 567 500  
 FT 568 500  
 FT 569 500  
 FT 570 500  
 FT 571 500  
 FT 572 500  
 FT 573 500  
 FT 574 500  
 FT 575 500  
 FT 576 500  
 FT 577 500  
 FT 578 500  
 FT 579 500  
 FT 580 500  
 FT 581 500  
 FT 582 500  
 FT 583 500  
 FT 584 500  
 FT 585 500  
 FT 586 500  
 FT 587 500  
 FT 588 500  
 FT 589 500  
 FT 590 500  
 FT 591 500  
 FT 592 500  
 FT 593 500  
 FT 594 500  
 FT 595 500  
 FT 596 500  
 FT 597 500  
 FT 598 500  
 FT 599 500  
 FT 600 500  
 FT 601 500  
 FT 602 500  
 FT 603 500  
 FT 604 500  
 FT 605 500  
 FT 606 500  
 FT 607 500  
 FT 608 500  
 FT 609 500  
 FT 610 500  
 FT 611 500  
 FT 612 500  
 FT 613 500  
 FT 614 500  
 FT 615 500  
 FT 616 500  
 FT 617 500  
 FT 618 500  
 FT 619 500  
 FT 620 500  
 FT 621 500  
 FT 622 500  
 FT 623 500  
 FT 624 500  
 FT 625 500  
 FT 626 500  
 FT 627 500  
 FT 628 500  
 FT 629 500  
 FT 630 500  
 FT 631 500  
 FT 632 500  
 FT 633 500  
 FT 634 500  
 FT 635 500  
 FT 636 500  
 FT 637 500  
 FT 638 500  
 FT 639 500  
 FT 640 500  
 FT 641 500  
 FT 642 500  
 FT 643 500  
 FT 644 500  
 FT 645 500  
 FT 646 500  
 FT 647 500  
 FT 648 500  
 FT 649 500  
 FT 650 500  
 FT 651 500  
 FT 652 500  
 FT 653 500  
 FT 654 500  
 FT 655 500  
 FT 656 500  
 FT 657 500  
 FT 658 500  
 FT 659 500  
 FT 660 500  
 FT 661 500  
 FT 662 500  
 FT 663 500  
 FT 664 500  
 FT 665 500  
 FT 666 500  
 FT 667 500  
 FT 668 500  
 FT 669 500  
 FT 670 500  
 FT 671 500  
 FT 672 500  
 FT 673 500  
 FT 674 500  
 FT 675 500  
 FT 676 500  
 FT 677 500  
 FT 678 500  
 FT 679 500  
 FT 680 500  
 FT 681 500  
 FT 682 500  
 FT 683 500  
 FT 684 500  
 FT 685 500  
 FT 686 500  
 FT 687 500  
 FT 688 500  
 FT 689 500  
 FT 690 500  
 FT 691 500  
 FT 692 500  
 FT 693 500  
 FT 694 500  
 FT 695 500  
 FT 696 500  
 FT 697 500  
 FT 698 500  
 FT 699 500  
 FT 700 500  
 FT 701 500  
 FT 702 500  
 FT 703 500  
 FT 704 500  
 FT 705 500  
 FT 706 500  
 FT 707 500  
 FT 708 500  
 FT 709 500  
 FT 710 500  
 FT 711 500  
 FT 712 500  
 FT 713 500  
 FT 714 500  
 FT 715 500  
 FT 716 500  
 FT 717 500  
 FT 718 500  
 FT 719 500  
 FT 720 500  
 FT 721 500  
 FT 722 500  
 FT 723 500  
 FT 724 500  
 FT 725 500  
 FT 726 500  
 FT 727 500  
 FT 728 500  
 FT 729 500  
 FT 730 500  
 FT 731 500  
 FT 732 500  
 FT 733 500  
 FT 734 500  
 FT 735 500  
 FT 736 500  
 FT 737 500  
 FT 738 500  
 FT 739 500  
 FT 740 500  
 FT 741 500  
 FT 742 500  
 FT 743 500  
 FT 744 500  
 FT 745 500  
 FT 746 500  
 FT 747 500  
 FT 748 500  
 FT 749 500  
 FT 750 500  
 FT 751 500  
 FT 752 500  
 FT 753 500  
 FT 754 500  
 FT 755 500  
 FT 756 500  
 FT 757 500  
 FT 758 500  
 FT 759 500  
 FT 760 500  
 FT 761 500  
 FT 762 500  
 FT 763 500  
 FT 764 500  
 FT 765 500  
 FT 766 500  
 FT 767 500  
 FT 768 500  
 FT 769 500  
 FT 770 500  
 FT 771 500  
 FT 772 500  
 FT 773 500  
 FT 774 500  
 FT 775 500  
 FT 776 500  
 FT 777 500  
 FT 778 500  
 FT 779 500  
 FT 780 500  
 FT 781 500  
 FT 782 500  
 FT 783 500  
 FT 784 500  
 FT 785 500  
 FT 786 500  
 FT 787 500  
 FT 788 500  
 FT 789 500  
 FT 790 500  
 FT 791 500  
 FT 792 500  
 FT 793 500  
 FT 794 500  
 FT 795 500  
 FT 796 500  
 FT 797 500  
 FT 798 500  
 FT 799 500  
 FT 800 500  
 FT 801 500  
 FT 802 500  
 FT 803 500  
 FT 804 500  
 FT 805 500  
 FT 806 500  
 FT 807 500  
 FT 808 500  
 FT 809 500  
 FT 810 500  
 FT 811 500  
 FT 812 500  
 FT 813 500  
 FT 814 500  
 FT 815 500  
 FT 816 500  
 FT 817 500  
 FT 818 500  
 FT 819 500  
 FT 820 500  
 FT 821 500  
 FT 822 500  
 FT 823 500  
 FT 824 500  
 FT 825 500  
 FT 826 500  
 FT 827 500  
 FT 828 500  
 FT 829 500  
 FT 830 500  
 FT 831 500  
 FT 832 500  
 FT 833 500  
 FT 834 500  
 FT 835 500  
 FT 836 500  
 FT 837 500  
 FT 838 500  
 FT 839 500  
 FT 840 500  
 FT 841 500  
 FT 842 500  
 FT 843 500  
 FT 844 500  
 FT 845 500  
 FT 846 500  
 FT 847 500  
 FT 848 500  
 FT 849 500  
 FT 850 500  
 FT 851 500  
 FT 852 500  
 FT 853 500  
 FT 854 500  
 FT 855 500





ON	NCR1_TextID=4081;
RN	[1]
RP	SEQUENCE FROM N.A.
FX	MEDLINE=90239044; PubMed=2185473;
RA	Cottelman S., Dalrymple C., Pichersky E., Cartington M., Hobbs M.,
RA	Mallack J.S., Saliques B., Kurimatsu H., Shiroza T., Foster T.,
RA	Clark W.P., Ross B., Saliques C.V., Mauritz M.R.;
RT	Abstract In Proteins Involved in the regulatory subunit for the Clp ATP-dependent
RT	protease in prokaryotes. J Biol Chem 266(19):11990-11995 (1991).
RL	Proc. Natl. Acad. Sci. U.S.A. 87:3513-3517(1990).
CC	-1- FUNCTION: MAY INTERACT WITH A CLP-LIKE PROTEASE INVOLVED IN
CC	DENATURATION OF DENATURED PROTEINS IN THE CHLOROPLAST.
CC	-1- SUBCELLULAR LOCATION: CHLOROPLAST.
CC	-1- SIMILARITY: BELONGS TO THE CLPA/CLEP FAMILY.
CC	This SwissProt entry is derived from a cDNA clone that was produced through a collaboration
CC	between the Swiss Institute Bioinformatics Institute. There are no restriction sites
CC	use by non-profit institutions is not removed as its content is in no way
CC	modified and this statement is as long as its removal is in no way
CC	entails requires a license agreement (see http://www.isb-sib.ch/announce/
CC	or send an email to licenses@isb.ch).
DR	EMBL: M2604; AAA3161.1; -
DR	PIR: B35605; B35605.1; -
DR	InterPro: IPRO03593; AAA.
DR	InterPro: IPRO01930; AAA_subfam.
DR	InterPro: IPRO01270; CLP_AB.
DR	InterPro: IPRO01943; VVR.
DR	Pfam: PF00004; AAA_2.
DR	Pfam: PF23134; VVR.
DR	PRINTS: PR01010; VVR.
DR	PROSITE: PS00870; CLPA1.1; 1.
DR	SMART: SM00382; AAA_2.
DR	PROSITE: PS00871; CLPA2.2; 1.
KW	Chaperone; ATP-binding; Repeat; Chloroplast; Transit peptide.
FT	TRANSIT ? 923
FT	CHAIN ? 923
FT	CDROPFAST (POTENTIAL).
FT	CLP-DEPENDENT CLP PROTEASE ATP-BINDING
FT	SUBUNIT CLPA HOMOLOG CDAB.
FT	DOMAIN 255 502
FT	DOMAIN 569 760
FT	DOMAIN 300 307
FT	NP_BIND 643 650
FT	ATP (POTENTIAL).
FT	ATP (POTENTIAL).
SQ	SEQUENCE 923 AA: 102241 MW: 26569DECEDE974B65 CRC64:
Qy	1 GCGTSG 6
Db	244 GCGTSG 249
Query Match	100.0%; Score 33; DB 1; Length 923;
Best Local Similarity	100.0%; Pred. No. 1 de+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
RESULT 17	
ID	CLPA.LYCES
AC	P31541;
STANDARD;	FRG; 926 AA.
DY	01-JUL-1993 (Rel. 26, Created)
DY	01-JUL-1993 (Rel. 26, Last sequence update)
DY	01-FEB-1996 (Rel. 37, Last annotation update)
DT	PREPENDANT CLP PROT33
DT	PREPENDANT CLP PROT33
GN	CDAA.
OC	Lycopersicon esculentum (Tomato).
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
OC	Asteridae; eusterids I; Solanales; Solanaceae; Solanum.
OC	NCBI_TaxId=4081;
RP	SEQUENCE FROM N.A.
FX	MEDLINE=90239044; PubMed=2185473;



RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakae S., Nakamura Y., Nishimoto H.,  
 RA Nishimura S., Saito M., Sempel G., Stieh Y., Sivasubram S.,  
 RA Taniguchi H., Tsubaki T., Takeda Y., Takemoto K., Uehara K., Mada C.,  
 RA Yamagata S., Horizuchi T., Takeda Y., Takemoto K., Uehara K., Mada C.,  
 RT \*Construction of a contiguous 874-bp sequence of the Escherichia coli  
 RT K-12 genome corresponding to 50.0-68.8 mln on the linkage map and  
 RT analysis of its sequence features.\*  
 RT DNA Res. 4:91-113(1997).  
 RP SEQUENCE FROM N.A.  
 RX STRAIN:01-57:HT / EDL933 / ATCC 700927;  
 RX MEDLINE:21074935; PubMed:11206551;  
 RA Perna N.T., Plunkett G., III, Burland V., Mau B., Glasner J.D.,  
 RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Klapavick H.A.,  
 RA Fostel G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
 RA Apodaca E., Davis M.W., Lam A., Dimalanta E.T., Potamianos K.,  
 RA Welch R.A., Blattner F.R., Ito J., Ten G., Schwartz D.C.,  
 RT \*Genome sequence of enterohemorrhagic Escherichia coli O157:HT.\*  
 RT Nature 409:529-533(2001).  
 RL Sequence from N.A.  
 RP MEDLINE:211571; PubMed:1058952;  
 RX MEDLINE:211571; PubMed:1058952;  
 RA Hayashi T., Makino K., Ohnishi K., Kurikawa K., Taki K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tanaka T.,  
 RA Iida T., Takami H., Honda T., Sasakiwa C., Ogasawara N., Tsunaga T.,  
 RA Kohara S., Shiba T., Hattori M., Shigaawa H.;  
 RT \*Complete genome sequence of enterohemorrhagic Escherichia coli  
 RT O157:H7 and genomic comparison with a laboratory strain K-12.\*  
 RL DNA Res. 4:120-127(1997).  
 CC -1- FUNCTION. NOT KNOWN. PROBABLY INVOLVED IN RIBONUCLEOTIDE REDUCTASE  
 CC FUNCTION.  
 CC -1- SIMILARITY: BELONGS TO THE NBD1 FAMILY.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sdb.ch/announce/>  
 CC or send an email to [license@1sb-sdb.ch](mailto:license@1sb-sdb.ch)).  
 CC  
 CC EMBL: X75787; CNA56185.1;  
 CC EMBL: X75787; CNA56185.1;  
 CC EMBL: D5089; BAE5521.1;  
 CC EMBL: D5089; BAE5521.1;  
 CC EMBL: AE005496; AAG57784.1;  
 CC EMBL: AE005496; AAG57784.1;  
 CC EMBL: EC13287; nrd1.  
 CC EMBL: EC13287; nrd1.  
 CC InterPro: IPR003707; Nrd1.  
 CC Pfam: PF02555; Nrd1; 1.  
 CC Complete proteome.  
 CC Sequence 136 AA; 15340 MW; 2834DB08D18BCBB CRC64;  
 CC  
 CC Query Match 90.9%; Score 30; DB 1; Length 136;  
 CC Best Local Similarity 83.3%; Pred. No. 87;  
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC 1 GGGTSG 6  
 CC  
 CC 51 GGGRGA 56  
 CC  
 CC RESULT 20  
 CC TUFFEYFAST STANDARD; PRT; 245 AA.  
 CC  
 CC 01-FEB-1996 (Rel. 33, Created)  
 CC 01-FEB-1996 (Rel. 33, Last sequence update)  
 CC 01-NOV-1997 (Rel. 35, Last annotation update)

DE HYPOTHETICAL 26.9 KDA PROTEIN IN EBN1-PE8 INTERGENIC REGION.  
 CS XJ1055W OR J1146  
 CS Saccharomyces cerevisiae (Baker's yeast).  
 CC Eukaryota: Fungi; Ascomycota: Saccharomycotina: Saccharomycetes;  
 CC Saccharomycetales: Saccharomycetaceae: Saccharomycos.  
 CC NCBI\_TaxID=4932.  
 CC  
 CC SEQUENCE FROM N.A.  
 RP  
 RA Pohl T.M., Allmönig G.;  
 RT Submitted (SEP-1995) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: TO R.FASCIANS HYPOTHETICAL 21.1 KDA PROTEIN IN  
 CC FASCINATION LOCUS (AC P46378).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sdb.ch/announce/>  
 CC or send an email to [license@1sb-sdb.ch](mailto:license@1sb-sdb.ch)).  
 CC  
 CC EMBL: Z49330; C489346.1;  
 CC SCD: S0003591; XJ1055W  
 CC Hypothetical protein: ATP-binding.  
 RT NP\_BIND 13  
 RT SEQUENCE 245 AA; 26864 MW; 0C49EFCF42E7CA2 CRC64;  
 CC  
 CC Query Match 90.9%; Score 30; DB 1; Length 245;  
 CC Best Local Similarity 83.3%; Pred. No. 136+02;  
 CC Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 CC  
 CC 1 GGGTSG 6  
 CC  
 CC 56 GGGRGA 61  
 CC  
 CC RESULT 21  
 CC IF31\_HUMAN STANDARD; PRT; 258 AA.  
 CC  
 CC 0758823;  
 CC 30-MAY-2000 (Rel. 38, Created)  
 CC 20-MAY-2000 (Rel. 38, Last sequence update)  
 CC 20-MAY-2000 (Rel. 38, Last annotation update)  
 CC Eukaryotic TRANSLATION INITIATION FACTOR 3 SUBUNIT 1 (EIF-3 ALPHA)  
 DE (EIF3 P35)  
 GN EIF3.1.  
 OS Homo sapiens (Human).  
 CC Eukaryota: Chordata: Vertebrata: Euteleostomi;  
 CC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
 CC NCBI\_TaxID=9606;  
 CC  
 CC SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RX MEDLINE:99041954; PubMed:9822653.  
 RA Block K.L., Vornlocher H.-P., Herzberg J.W.B.;  
 RT \*Characterization of cDNAs encoding the p44 and p35 subunits of human  
 RT E. Biol. Chem. 273:31901-31908(1998).  
 CC -1- METHIONYL-TRNA AND RNA  
 CC  
 CC -1- SUBUNIT: EIF-3 IS COMPOSED OF AT LEAST 10 DIFFERENT SUBUNITS.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1sb-sdb.ch/announce/>  
 CC or send an email to [license@1sb-sdb.ch](mailto:license@1sb-sdb.ch)).  
 CC  
 CC EMBL: U97670; AAC78729.1;  
 CC MIM: 603910;  
 CC Initiation factor; Protein biosynthesis.

FT DOMAIN 2 8 POLY-A1A  
 FT DOMAIN 29 32 POLY-GLY  
 FT DOMAIN 53 57 POLY-ASP  
 FT DOMAIN 218 224 POLY-LYS  
 SEQUENCE 258 AA: 28990 MW: F01E75741473058B CRC64;

Query Match 90.94; Score 30; DB 1; Length 258;

Best Local Similarity 83.34; Pred. No. 1-be-02; Mismatches 0; Gaps 0;

Matches 5; Conservative 1; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
 DB 30 GCGTSG 35

## RESULT 22

EUT1\_ECOLI STANDARD: PRT; 278 AA.  
 AC P77277;

DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ETHANOLAMINE UTILIZATION PROTEIN EUT1

GN EUT1 OR B2454.  
 OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia

RN [1] TaxID=562;  
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MO1655;  
 RX MEDLINE=97426617; PubMed=9279503;

RA Blatner F.R., Plunkett G., III, Bloch C.A., Berna N.T., Burland V.,  
 Riley M., Rhee V., Colado-Vides J., Glander J.D., Glick C.K., Haynes G.F.,

RA Kelley R., Klenk A.P., Knapik G., Krumholz L., Ladendorfer J., Lapidus  
 R.A. M.B., Shao Y.; N.M., Karpelitz H.A., Goclen N.A., Rose D.J.,

RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).

RN [3]  
 RP SEQUENCE FROM N.A.

RC STRAIN=K12 / MO1655;  
 RX MEDLINE=97426617; PubMed=9205637;

RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,  
 Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsushashi N.,

RA Mitsuuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,  
 RA Ohshima T., Oyama S., Saito N., Sempel G., Setoh Y., Sivasubramanian S.,

RA Tegenfl H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata H., Yoshida T., Yoshida T., Yoshida T., Yoshida T., Yoshida T.,

RA - K12 genome corresponding to 50,066,8 mln on the linkage map and  
 RT analysis of its sequence features."

RL DNA Res. 4:91-113(1997).  
 CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the European Bioinformatics Institute and the Swiss Institute of Bioinformatics.

CC The European Bioinformatics Institute and the Swiss Institute of Bioinformatics  
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL: S0000332; AAC01507.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

DR EMBL: D90874; AAC6322.1;

OY 1 GCGTSG 6  
 DB 145 GCGTSG 150

## RESULT 23

EUT1\_SALTY STANDARD: PRT; 279 AA.  
 AC P41794;

DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)

DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE ETHANOLAMINE UTILIZATION PROTEIN EUT1

GN EUT1  
 OS Salmonella typhimurium.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Salmonella

RN [1] TaxID=602;  
 RP SEQUENCE FROM N.A.

RC STRAIN=ATCC 14028S;  
 RX MEDLINE=95173114; PubMed=766611;

RA Stojiljkovic I., Baewer A.J., Heffron F.,  
 RA "Ethanolamine utilization in Salmonella typhimurium: nucleotide

RA sequence, protein expression, and mutational analysis of the ccmA  
 RT ccmA gene, eutC gene, and eutD gene."

RL J. Bacteriol. 177:1557-1560(1995).  
 RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=LT2;  
 RX MEDLINE=95395039; PubMed=10464203;

RA Kofoid B.C., Rappleye C.A., Stojiljkovic I., Roth J.R.,  
 RA "The gene ethanologin of Salmonella typhimurium

RA encodes the ethanologin of carboxysome shell proteins."  
 RL J. Bacteriol. 181:5317-5329(1999)

CC -1- PATHWAY: ETHANOLAMINE UTILIZATION.

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the European Bioinformatics

CC The European Bioinformatics Institute and the Swiss Institute of Bioinformatics  
 use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL: U08500; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;

DR EMBL: AF093105; AA080210.1;





Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
 OY 1 CGCTSG 6  
 DB 233 CGCTG 238

RESULT 29  
 DHA\_BAGCT STANDARD: PRT: 372 AA.  
 AC P14553  
 DT 01-AUG-1990 (Rel. 15, Created)  
 DT 01-AUG-1990 (Rel. 15, Last sequence update)  
 DT 01-OCT-1994 (Rel. 30, Last annotation update)  
 DE ALANINE DEHYDROGENASE (EC 1.4.1.1).  
 OS Bacillus stearothermophilus.  
 CC Bacillus, Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Geobacillus.  
 CX MOLTEXT:ID=1422.  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RC STRAIN-IFO 12550;  
 RA MEDLINE=9054112; PubMed=2340274;  
 RA Kuroda S., Tanizawa K., Sakamoto Y., Tanaka H., Soda K.:  
 R1 Alantate dehydrogenases from two *Bacillus* species with distinct  
 R2 determination of molecular cloning, DNA and protein sequence  
 R3 MD(P)(+)-dependent dehydrogenases.  
 R4 Biochemistry 29:1009-1015(1990)  
 CC -1- FUNCTION: THIS ENZYME IS A KEY FACTOR IN THE ASSIMILATION OF  
 CC L-ALANINE AS AN ENERGY SOURCE THROUGH THE TRICARBOXYLIC ACID  
 CC CYCLE DURING SPOULATION.  
 CC -1- SUBUNIT ACTIVITY: L-ALANINE + H(2)O + NAD(+) = PYRUVATE +  
 CC NH(3)  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC.  
 CC -1- MISCELLANEOUS: THE B. STEAROTHERMOPHILUS ENZYME RETAINED ABOUT 50%  
 CC OF ITS INITIAL ACTIVITY WHEN HEATED AT 85 DEGREES CELSIUS FOR 5  
 CC MIN AT PH 7.2. WHEREAS THE B. SPHERICUS ENZYME LOST THE SAME  
 CC -1- SIMILARITY: WHEN HEATED AT ONLY 65 DEGREES CELSIUS FOR 5 MIN.  
 CC WITH PYRIDINE NUCLEOTIDE TRANSFERASES.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 CC EMBL: M33299; AAA2211.1;  
 DR PIR: B34261; B34261.  
 DR InterPro: IPRO00598; Aladh\_PNT.  
 DR Pfam: PF01262; Aladh\_PNT.1; Aladh\_PNT.2.  
 DR PROSITE: PS00836; Aladh\_PNT.1; 1.  
 DR PROSITE: PS00837; Aladh\_PNT.2; 1.  
 KW Oxidoreductase; NAD; Sporulation.  
 FT ACT\_SITE 96  
 FT POTENTIAL 96  
 FT NC\_BIND 170  
 FT SEQUENCE 372 AA; 39694 MW; 99A8E851ED1F73A CRC64;

Query Match 90.9%; Score 30; DB 1; Length 372;  
 Best Local Similarity 83.3%; Pred.No. 2,3e+02;  
 Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 CGCTSG 6  
 DB 175 GCGTGA 180

RESULT 30  
 LKXA\_MESAU STANDARD: PRT: 382 AA.  
 ID O04650.  
 DT 01-JUN-1994 (Rel. 28, Created)  
 DT 01-JUN-1994 (Rel. 28, Last sequence update)  
 DT 01-JUN-1994 (Rel. 28, Last annotation update)  
 DE LIM/HOMEOBOX PROTEIN LKXA (LIM-1) (LIM-HOMEOBOX PROTEIN 1).  
 GN LKXA OR LKX1  
 OS Mesocricetus auratus (Golden hamster).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Cricetinae;  
 CC Mesocricetus  
 CC Mesocricetus  
 CX MOLTEXT:ID=10036;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Pancreas;  
 RA MEDLINE=9305335; PubMed=1388758;  
 RA German M.S., Wang J., Chadwick R.B., Rutter W.J.:  
 R1 Synergistic activation of the insulin gene by a LIM-homeo domain  
 R2 protein and a basic helix-loop-helix protein: building a functional  
 R3 protein in an enhancer complex.  
 R4 Science 265:1232-1235(1994)  
 CC -1- FUNCTION: ACTS AS A TRANS-CELL-SPECIFIC TRANSCRIPTIONAL  
 CC FLAT ELEMENT WHICH IS A BETA-CELL-SPECIFIC TRANSCRIPTIONAL  
 CC ITS LIM-CONTAINING AMINO TERMINUS WITH SH-PAN, A HELIX-LOOP-HELIX  
 CC PROTEIN.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.  
 CC -1- SUBUNIT ACTIVITY: TO OTHER HOMEOBOX DOMAINS. BELONGS TO THE LIM  
 CC -1- SIMILARITY: CONTAINS 2 LIM DOMAINS. THE LIM DOMAIN BINDS 2 ZINC  
 CC IONS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC entitles requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 CC EMBL: J81406; CA537163.1;  
 DR PIR: J81406; CA537163.  
 DR HSP: P02836; LIMD.  
 DR TRANSFAC: T01958.  
 DR InterPro: IPRO01356; Homeobox.  
 DR InterPro: IPRO01781; LIM.  
 DR Pfam: PF00046; homeobox.1.  
 DR Pfam: PF00412; LIM.2.  
 DR SMART: SM00694; LIM.2.  
 DR SMART: SM00132; LIM.2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.1; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.2; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.3; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.4; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.5; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.6; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.7; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.8; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.9; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.10; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.11; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.12; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.13; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.14; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.15; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.16; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.17; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.18; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.19; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.20; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.21; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.22; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.23; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.24; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.25; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.26; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.27; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.28; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.29; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.30; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.31; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.32; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.33; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.34; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.35; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.36; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.37; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.38; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.39; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.40; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.41; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.42; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.43; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.44; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.45; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.46; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.47; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.48; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.49; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.50; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.51; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.52; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.53; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.54; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.55; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.56; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.57; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.58; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.59; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.60; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.61; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.62; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.63; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.64; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.65; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.66; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.67; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.68; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.69; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.70; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.71; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.72; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.73; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.74; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.75; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.76; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.77; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.78; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.79; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.80; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.81; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.82; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.83; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.84; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.85; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.86; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.87; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.88; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.89; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.90; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.91; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.92; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.93; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.94; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.95; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.96; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.97; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.98; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.99; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.100; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.101; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.102; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.103; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.104; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.105; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.106; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.107; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.108; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.109; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.110; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.111; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.112; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.113; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.114; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.115; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.116; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.117; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.118; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.119; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.120; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.121; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.122; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.123; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.124; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.125; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.126; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.127; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.128; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.129; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.130; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.131; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.132; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.133; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.134; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.135; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.136; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.137; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.138; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.139; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.140; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.141; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.142; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.143; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.144; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.145; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.146; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.147; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.148; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.149; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.150; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.151; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.152; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.153; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.154; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.155; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.156; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.157; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.158; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.159; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.160; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.161; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.162; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.163; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.164; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.165; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.166; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.167; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.168; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.169; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.170; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.171; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.172; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.173; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.174; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.175; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.176; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.177; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.178; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.179; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.180; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.181; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.182; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.183; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.184; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.185; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.186; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.187; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.188; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.189; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.190; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.191; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.192; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.193; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.194; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.195; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.196; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.197; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.198; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.199; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.200; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.201; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.202; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.203; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.204; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.205; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.206; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.207; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.208; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.209; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.210; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.211; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.212; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.213; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.214; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.215; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.216; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.217; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.218; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.219; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.220; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.221; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.222; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.223; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.224; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.225; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.226; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.227; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.228; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.229; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.230; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.231; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.232; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.233; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.234; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.235; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.236; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.237; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.238; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.239; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.240; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.241; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.242; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.243; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.244; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.245; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.246; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.247; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.248; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.249; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.250; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.251; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.252; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.253; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.254; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.255; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.256; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.257; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.258; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.259; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.260; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.261; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.262; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.263; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.264; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.265; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.266; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.267; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.268; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.269; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.270; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.271; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.272; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.273; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.274; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.275; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.276; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.277; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.278; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.279; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.280; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.281; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.282; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.283; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.284; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.285; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.286; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.287; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.288; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.289; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.290; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.291; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.292; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.293; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.294; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.295; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.296; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.297; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.298; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.299; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.300; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.301; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.302; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.303; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.304; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.305; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.306; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.307; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.308; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.309; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.310; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.311; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.312; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.313; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.314; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.315; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.316; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.317; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.318; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.319; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.320; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.321; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.322; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.323; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.324; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.325; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.326; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.327; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.328; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.329; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.330; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.331; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.332; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.333; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.334; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.335; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.336; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.337; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.338; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.339; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.340; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.341; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.342; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.343; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.344; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.345; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.346; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.347; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.348; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.349; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.350; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.351; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.352; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.353; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.354; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.355; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.356; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.357; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.358; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.359; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.360; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.361; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.362; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.363; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.364; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.365; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.366; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.367; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.368; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.369; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.370; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.371; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.372; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.373; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.374; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.375; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.376; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.377; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.378; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.379; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.380; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.381; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.382; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.383; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.384; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.385; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.386; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.387; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.388; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.389; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.390; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.391; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.392; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.393; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.394; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.395; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.396; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.397; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.398; 2.  
 DR PROSITE: PS00478; LIM\_DOMAIN.399; 2.  
 DR PROSITE





20-AUG-2001 (Rel. 40, Last annotation update)  
 DT UDP-N-ACETYLGLUCOSAMINE-N-ACETYLTRANSFERASE (PENTAPEPTIDE)  
 DE (EC 2.3.1.11) (UNDECAPRENTL-N-ACETYLGLUCOSAMINE TRANSFERASE  
 DE TRANSFERASE)  
 GN MORG OR M10914 OR M128268 02C.  
 OS Mycobacterium leprae.  
 CC Bacteria; Firmicutes; Actinobacteriia; Mycobacteriaceae; Mycobacterium.  
 CC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 CC Ref. taxid:1769.  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-TN:  
 RX MEDLINE-21128732: PubMed-11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Holtz S., Horsley J., Jones S., Kaul R., Maclean J., Moule S.,  
 RA Murphy L., Oliver J., Quail M.A., Rajagopal K., Rutherford K.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RA Massive genome decay in the leprosy bacillus. ?  
 RT PubMed:11234002(2001).  
 CC -1- FUNCTION: CELL WALL FORMATION. CATALYZES THE TRANSFER OF A GLUCNA  
 CC SUBUNIT ON UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE I) TO FORM UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE II) (BY SIMILARITY).  
 CC -1- PATHWAY: LAST STEP OF PEPTIDOGLYCAN BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE MORG FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL database  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC or send an email to license@sb-sib.ch).  
 CC EMBL: AL022602; CA18668.1; -  
 DR Lepidoptera: M10914; -  
 DR Transferrase: Glycosyltransferase; Cell division; Cell wall; Membrane;  
 KW Peptidoglycan synthesis; Complete proteome.  
 SQ SOURCE 407 AA: 42422 MW: 67518730935399 CR664.  
 Query Match 90.9%: Score 30; DB 1; Length 407;  
 Best Local Similarity 83.3%: Pred. No.2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 Db 35 GGGTNC 40

20-AUG-2001 (Rel. 40, Last annotation update)  
 DT UDP-N-ACETYLGLUCOSAMINE-N-ACETYLTRANSFERASE (PENTAPEPTIDE)  
 DE (EC 2.3.1.11) (UNDECAPRENTL-N-ACETYLGLUCOSAMINE TRANSFERASE  
 DE TRANSFERASE)  
 GN MORG OR M10914 OR M128268 02C.  
 OS Mycobacterium leprae.  
 CC Bacteria; Firmicutes; Actinobacteriia; Mycobacteriaceae; Mycobacterium.  
 CC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 CC Ref. taxid:1769.  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-TN:  
 RX MEDLINE-21128732: PubMed-11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Holtz S., Horsley J., Jones S., Kaul R., Maclean J., Moule S.,  
 RA Murphy L., Oliver J., Quail M.A., Rajagopal K., Rutherford K.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RA Massive genome decay in the leprosy bacillus. ?  
 RT PubMed:11234002(2001).  
 CC -1- FUNCTION: CELL WALL FORMATION. CATALYZES THE TRANSFER OF A GLUCNA  
 CC SUBUNIT ON UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE I) TO FORM UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE II) (BY SIMILARITY).  
 CC -1- PATHWAY: LAST STEP OF PEPTIDOGLYCAN BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE MORG FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL database  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC or send an email to license@sb-sib.ch).  
 CC EMBL: 295388; CAB08640.1; -  
 DR EMBL: A6007068; AAA64496.1; -  
 DR Transferrase: Glycosyltransferase; Cell division; Cell wall; Membrane;  
 KW Peptidoglycan synthesis; Complete proteome.  
 SQ SOURCE 410 AA: 41860 MW: 96502496356308 CR664.  
 Query Match 90.9%: Score 30; DB 1; Length 410;  
 Best Local Similarity 83.3%: Pred. No.2.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 Db 42 GGGTNC 47

20-AUG-2001 (Rel. 40, Last annotation update)  
 DT UDP-N-ACETYLGLUCOSAMINE-N-ACETYLTRANSFERASE (PENTAPEPTIDE)  
 DE (EC 2.3.1.11) (UNDECAPRENTL-N-ACETYLGLUCOSAMINE TRANSFERASE  
 DE TRANSFERASE)  
 GN MORG OR M10914 OR M128268 02C.  
 OS Mycobacterium leprae.  
 CC Bacteria; Firmicutes; Actinobacteriia; Mycobacteriaceae; Mycobacterium.  
 CC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 CC Ref. taxid:1769.  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-TN:  
 RX MEDLINE-21128732: PubMed-11234002;  
 RA Cole S.T., Eigmeier K., Parkhill J., James K.D., Thomson N.R.,  
 RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
 RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Holtz S., Horsley J., Jones S., Kaul R., Maclean J., Moule S.,  
 RA Murphy L., Oliver J., Quail M.A., Rajagopal K., Rutherford K.,  
 RA Rutter S., Seeger K., Simon S., Simmonds M., Skellon J., Squares R.,  
 RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
 RA Barrett B.G.;  
 RA Massive genome decay in the leprosy bacillus. ?  
 RT PubMed:11234002(2001).  
 CC -1- FUNCTION: CELL WALL FORMATION. CATALYZES THE TRANSFER OF A GLUCNA  
 CC SUBUNIT ON UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE I) TO FORM UNDECAPRENTL-PROPHOSPHORYLMURAMIC ACID (LIPID  
 CC INTERMEDIATE II) (BY SIMILARITY).  
 CC -1- PATHWAY: LAST STEP OF PEPTIDOGLYCAN BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: MEMBRANE ASSOCIATED (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE MORG FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL database  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC or send an email to license@sb-sib.ch).  
 CC EMBL: 295388; CAB08640.1; -  
 DR EMBL: A6007068; AAA64496.1; -  
 DR Transferrase: Glycosyltransferase; Cell division; Cell wall; Membrane;  
 KW Peptidoglycan synthesis; Complete proteome.  
 SQ SOURCE 410 AA: 41860 MW: 96502496356308 CR664.  
 Query Match 90.9%: Score 30; DB 1; Length 410;  
 Best Local Similarity 83.3%: Pred. No.2.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 Db 42 GGGTNC 47

NCBI\_TaxId=1773;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV.  
 RA MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eiglmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Davies K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies K., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Oliver S., Osborne J., Quail M.A., Richardson M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skellern S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead P., Barrett B.G.;  
 RA "Deciphering the biology of *Mycobacterium tuberculosis* from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gaitan M.L., Holt D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Esmailova M.D., Salzberg S.L.,  
 RA Delcher A.L., Utterback T., Weidman J., Gill J., Mikula A.,  
 RA Bernal W.;  
 RA "Genomic comparison of *Mycobacterium tuberculosis* clinical and  
 RT laboratory strains."  
 RL Submitted (APR-2001) to the EMBL/GenBank/CDDB databases.  
 CC -1- SIMILARITY: BELONGS TO THE MYCOBACTERIAL PE FAMILY. PERS  
 SUBFAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration  
 CC between the European Bioinformatics Institute and the EMBL collaboration  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: A021867; CA11384.1; ALT\_LINKT.  
 CC EMBL: A0206991; A040353.1; ALT\_LINKT.  
 CC TIGR: M11097.  
 DR Tubercule: Rv1068c; -  
 DR InterPro: IPR02052; Eggshell.  
 DR InterPro: IPR000084; PE.  
 DR Pfam: PF00934; PE: 1.  
 DR Pfam: PF01423; Eggshell.  
 DR Pfam: PF01423; Eggshell.  
 KM Hypothetical protein: Complete proteome.  
 FT CONFLICT 218 218 T -> S (IN REF. 2).  
 FT 235 235 G -> GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG (IN REF.  
 FT 2) GGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG (IN REF.  
 FT  
 FT SEQUENCE 463 AA: 39305 MW: CF96667E9393952 CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 463;  
 Best Local Similarity 83.3%; Pred. No. 2.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTSG 6  
 DB 286 GGGTAG 291  
 RESULT 36  
 MDL3\_PUUSE STANDARD: PRT: 573 AA.  
 ID 050048  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE (R)-MANDELONITRILE LYASE ISOFORM 3 PRECURSOR (EC 4.1.2.10)  
 DE (HYDROXYNITRILE LYASE 3) ((R)-OXANITRILE 3).

MDL3.  
 GN Prunus serotina (black cherry).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Euphorbiales; Euphorbiaceae; Euphorbia; Core eudicots; Rosales;  
 OC Eucroside I; Rosales; Rosaceae; Prunus.  
 ON NCBI\_TaxId=21207.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98076189; PubMed=9414550;  
 RA Hsu Z., Boulton J.E.;  
 RA "Cloning, organization, and preliminary promoter analysis  
 RT of a black cherry (Prunus serotina) mandelonitrile lyase gene."  
 RL Plant Physiol. 115:1359-1369(1997).  
 CC -1- FUNCTION: INVOLVED IN CYANOLENIN, THE RELEASE OF HCN FROM INJURED  
 CC TISSUES. CATALYSES THE STEREOSPECIFIC ADDITION OF HCN TO A VARIETY  
 CC OF ALDEHYDES IN VITRO. IS A MAJOR SEED CONSTITUENT, AND COULD HAVE  
 CC THE ADDITIONAL ROLE OF A STORAGE FORM FOR REDUCED NITROGEN.  
 CC -1- COFACTOR AND COWORD: CYANIDE + BENZALDEHYDE.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: PRIMARILY FOUND WITHIN PROTEIN BODIES OF THE  
 CC CYTOSOL/PARENCHYMA CELLS WITH LESSER AMOUNTS WITHIN THE  
 CC PROCAMBIUM (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.  
 CC  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration  
 CC between the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announcements>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL: U51567; AA96782.1; -  
 CC EMBL: AF013161; AA86774.1; -  
 DR InterPro: IPR001072; GMC\_Oxred.  
 DR ProSite: PS00624; GMC\_Oxred\_2; 1.  
 DR Pfam: PF00732; GMC\_Oxred.  
 KM Euker: Glycoprotein; FMD; Flavonoid; Signal: Multigene family.  
 FT CHAIN 28 573  
 FT NP\_BIND 55 82  
 FT CARBOHYD 30 30 FAD (ADP PART) (PROBABLE).  
 FT CARBOHYD 75 75 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 145 145 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 150 150 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 159 159 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 218 218 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 252 252 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 267 267 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 309 309 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 380 380 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 400 400 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 420 420 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 457 457 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 FT CARBOHYD 467 467 N-LINKED (GLNLCNC - - -) (POTENTIAL).  
 SO SEQUENCE 573 AA: 62179 MW: DB1BC8B9D37800 CRC64;  
 Query Match 90.9%; Score 30; DB 1; Length 573;  
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGTSG 6  
 DB 60 GGGTAG 65  
 RESULT 37  
 MDL2\_PUUSE STANDARD: PRT: 576 AA.  
 ID 050048  
 DT 15-DEC-1998 (Rel. 37, Created)





Search completed: February 4, 2002, 08:01:10  
Job Name: 34 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:00:16 ; Search time 35.81 Seconds

(without alignments)  
24,508 Million cell updates/sec

Title: us-09-642-660-10-

Sequence: 1 GCGTSC 6

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 473505 seqs, 146272329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database :

1: sptRBM1.17.\*  
2: sptRBM1.17.\*  
3: sptRBM1.17.\*  
4: sptRBM1.17.\*  
5: sptRBM1.17.\*  
6: sptRBM1.17.\*  
7: sptRBM1.17.\*  
8: sptRBM1.17.\*  
9: sptRBM1.17.\*  
10: sptRBM1.17.\*  
11: sptRBM1.17.\*  
12: sptRBM1.17.\*  
13: sptRBM1.17.\*  
14: sptRBM1.17.\*

Prod. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Match	Length	DB	ID	Description
1	33	100.0	218	10	09J02	09J02	arabidopsis
2	33	100.0	221	12	09J02	09J02	arabidopsis
3	33	100.0	224	3	09H73	09H73	plumose
4	33	100.0	251	2	09EV39	09EV39	strepococc
5	33	100.0	256	4	09N70	09N70	homo sapien
6	33	100.0	256	4	09N75	09N75	homo sapien
7	33	100.0	273	5	09J10	09J10	drosophila
8	33	100.0	273	5	09J10	09J10	drosophila
9	33	100.0	289	10	09H40	09H40	caenorhabd
10	33	100.0	303	10	09JY4	09JY4	arabidopsis
11	33	100.0	315	5	09Y10	09Y10	amblyoma
12	33	100.0	427	8	09G43	09G43	homo sapien
13	33	100.0	444	8	09Y13	09Y13	pseudocier
14	33	100.0	444	8	09Y13	09Y13	selaginella
15	33	100.0	451	8	09I61	09I61	lophocoe
16	33	100.0	451	8	09I61	09I61	lophocoe
17	33	100.0	468	8	09G48	09G48	desmodia
18	33	100.0	468	8	09G48	09G48	complanata
19	33	100.0	470	4	09H10	09H10	homo sapien

20	33	100.0	475	8	09Y74	09Y74	rhynchosteg
21	33	100.0	475	8	09Y74	09Y74	rhynchosteg
22	33	100.0	483	3	09B17	09B17	homo sapien
23	33	100.0	484	5	09V59	09V59	drosophila
24	33	100.0	503	2	09A42	09A42	caulobacter
25	33	100.0	504	2	09A42	09A42	caulobacter
26	33	100.0	504	2	09A42	09A42	caulobacter
27	33	100.0	504	2	09A42	09A42	caulobacter
28	33	100.0	504	2	09A42	09A42	caulobacter
29	33	100.0	504	2	09A42	09A42	caulobacter
30	33	100.0	504	2	09A42	09A42	caulobacter
31	33	100.0	504	2	09A42	09A42	caulobacter
32	33	100.0	504	2	09A42	09A42	caulobacter
33	33	100.0	504	2	09A42	09A42	caulobacter
34	33	100.0	504	2	09A42	09A42	caulobacter
35	33	100.0	504	2	09A42	09A42	caulobacter
36	33	100.0	504	2	09A42	09A42	caulobacter
37	33	100.0	504	2	09A42	09A42	caulobacter
38	33	100.0	504	2	09A42	09A42	caulobacter
39	33	100.0	504	2	09A42	09A42	caulobacter
40	33	100.0	504	2	09A42	09A42	caulobacter
41	33	100.0	504	2	09A42	09A42	caulobacter
42	33	100.0	504	2	09A42	09A42	caulobacter
43	33	100.0	504	2	09A42	09A42	caulobacter
44	33	100.0	504	2	09A42	09A42	caulobacter
45	33	100.0	504	2	09A42	09A42	caulobacter
46	33	100.0	504	2	09A42	09A42	caulobacter
47	33	100.0	504	2	09A42	09A42	caulobacter
48	33	100.0	504	2	09A42	09A42	caulobacter
49	33	100.0	504	2	09A42	09A42	caulobacter
50	33	100.0	504	2	09A42	09A42	caulobacter
51	33	100.0	504	2	09A42	09A42	caulobacter
52	33	100.0	504	2	09A42	09A42	caulobacter
53	33	100.0	504	2	09A42	09A42	caulobacter
54	33	100.0	504	2	09A42	09A42	caulobacter
55	33	100.0	504	2	09A42	09A42	caulobacter
56	33	100.0	504	2	09A42	09A42	caulobacter
57	33	100.0	504	2	09A42	09A42	caulobacter
58	33	100.0	504	2	09A42	09A42	caulobacter
59	33	100.0	504	2	09A42	09A42	caulobacter
60	33	100.0	504	2	09A42	09A42	caulobacter
61	33	100.0	504	2	09A42	09A42	caulobacter
62	33	100.0	504	2	09A42	09A42	caulobacter
63	33	100.0	504	2	09A42	09A42	caulobacter
64	33	100.0	504	2	09A42	09A42	caulobacter
65	33	100.0	504	2	09A42	09A42	caulobacter
66	33	100.0	504	2	09A42	09A42	caulobacter
67	33	100.0	504	2	09A42	09A42	caulobacter
68	33	100.0	504	2	09A42	09A42	caulobacter
69	33	100.0	504	2	09A42	09A42	caulobacter
70	33	100.0	504	2	09A42	09A42	caulobacter
71	33	100.0	504	2	09A42	09A42	caulobacter
72	33	100.0	504	2	09A42	09A42	caulobacter
73	33	100.0	504	2	09A42	09A42	caulobacter
74	33	100.0	504	2	09A42	09A42	caulobacter
75	33	100.0	504	2	09A42	09A42	caulobacter
76	33	100.0	504	2	09A42	09A42	caulobacter
77	33	100.0	504	2	09A42	09A42	caulobacter
78	33	100.0	504	2	09A42	09A42	caulobacter
79	33	100.0	504	2	09A42	09A42	caulobacter
80	33	100.0	504	2	09A42	09A42	caulobacter
81	33	100.0	504	2	09A42	09A42	caulobacter
82	33	100.0	504	2	09A42	09A42	caulobacter
83	33	100.0	504	2	09A42	09A42	caulobacter
84	33	100.0	504	2	09A42	09A42	caulobacter
85	33	100.0	504	2	09A42	09A42	caulobacter
86	33	100.0	504	2	09A42	09A42	caulobacter
87	33	100.0	504	2	09A42	09A42	caulobacter
88	33	100.0	504	2	09A42	09A42	caulobacter
89	33	100.0	504	2	09A42	09A42	caulobacter
90	33	100.0	504	2	09A42	09A42	caulobacter
91	33	100.0	504	2	09A42	09A42	caulobacter
92	33	100.0	504	2	09A42	09A42	caulobacter





RC STRAIN-V288;  
 RC Bolken T.C., Franke C.A., Hruby D.E., Zeller G.O.;  
 RT Identification of New Intergenic and Intragenic Integration Sites for  
 RT Foreign Gene Expression in Recombinant *Streptococcus gordoni*  
 RT Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RL EMBL: AF210773; AAC93000.1;  
 DR InterPro: IPR003501; PFS.IIB.  
 DR Pfam: PF02302; PFS.IIB.1.  
 KM Transferase.  
 SO SOURCE 1  
 SEQUENCE 251 AA; 27174 MW; B8SCDFDFD8D33 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 251;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCGTSG 6  
 Db 157 GCGTSG 162

RESULT 5  
 G9WZT0 PRELIMINARY; PRT; 256 AA.  
 AC G9WZT0;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last annotation update)  
 DE PLUNC.  
 DE Homo sapiens (Human)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NX NCBI\_TaxId=9606;  
 RN [1]  
 RA Single C.D.:  
 RA Human Plunc, a novel upper airway enriched protein.\*;  
 RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF172993; AAF0860.1;  
 SO SEQUENCE 256 AA; 26712 MW; EDPL52FBD7EC555C CRC64;

Query Match 100.0%; Score 33; DB 4; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCGTSG 6  
 Db 80 GCGTSG 85

RESULT 6  
 G9NP55 PRELIMINARY; PRT; 256 AA.  
 ID G9NP55;  
 AC G9NP55;  
 DT 01-OCT-2000 (TRENBLrel. 15, Created)  
 DT 01-OCT-2000 (TRENBLrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE LUNG TISSUE ENRICHED PROTEIN (LUNG PROTEIN LOC51297).  
 DE EPITHELIUM ENRICHED PROTEIN (EAP5010.3) (LUNG PROTEIN LOC51297).  
 GN LUNG OR PLUNC OR BA9610.5.  
 OS Homo sapiens (Human)  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NX NCBI\_TaxId=9606;  
 RA Single C.D.:  
 RA Human Lung, a novel upper airway enriched protein.\*;  
 RP Tissue-Lung;  
 RA Matsubae T.;  
 RT \*Novel gene specifically expressed in lung.\*;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RA He Z.W., Yao K.T., Xu L.G., Ren C.P., Lan K., Xie L., Zhang L.,  
 RA Liu W.D., Zhou W., Wang L.;  
 RT Cloning a new gene related to nasopharyngeal carcinoma.\*;  
 RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20472055; PubMed=11018263;  
 RA Bingle C.D., Bingle L.;  
 RT Characterization of the human plunc gene, a gene product with an  
 RT upper airways and nasopharyngeal restricted expression pattern.\*;  
 RL Biochim. Biophys. Acta 1493:363-367(2000).  
 RP SEQUENCE FROM N.A.  
 RA Tracey A.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB024937; BA93633.1;  
 DR EMBL: AF158745; AAF82622.1;  
 DR EMBL: AF214562; AAG13653.1;  
 DR EMBL: AF121295; CAC03249.1;  
 SO SEQUENCE 256 AA; 26712 MW; EDPL52FBD7EC555C CRC64;

Query Match 100.0%; Score 33; DB 4; Length 256;  
 Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GCGTSG 6  
 Db 80 GCGTSG 85

RESULT 7  
 ID G9YTLO PRELIMINARY; PRT; 275 AA.  
 AC G9YTLO;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TRENBLrel. 17, Last annotation update)  
 DE CG11085 PROTEIN.  
 DE Drosophila melanogaster (Fruit fly)  
 OS Drosophila melanogaster (Fruit fly)  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephyridiidae; Drosophilidae; Drosophila.  
 NX NCBI\_TaxId=7227;  
 RN [1]  
 RA Single C.D.:  
 RA Human Plunc, a novel upper airway enriched protein.\*;  
 RP Tissue-Lung;  
 RA Matsubae T.;  
 RT \*Novel gene specifically expressed in lung.\*;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.G., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton R.C., Wortman J.R., Yandell M.D., Zhang O., Chen L.X.,  
 RA Brannon R.C., Rogers Y.H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
 RA Abul N.J., Doyle E.K., Baxter E.C., Nelt G., Nelson C.R., Miklos G.L.G.,  
 RA Avril J.F., Adams J.R., Baxendale J., Bayraktaroglu S., Galloway D.,  
 RA Bellow R.W., Bass A., Baxendale J., Bayraktaroglu S., Galloway D.,  
 RA Beeson K.Y., Benos P.V., Bonen B.J., Bhandal D., Boltskov S.,  
 RA Borkov D., Botchan M.R., Bouck J., Brooksfield P., Brothier P.,  
 RA Burtis J.K., Busam D.A., Butler G., Davenport L.B., Cantor A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke G., Davenport L.B., Davies P.,  
 RA Chervis J.M., Delecher A., Deng Z., Myers A.D., Dew T., Dietz S.M.,  
 RA DeLong E.M., DeLong E.M., DeLong E.M., DeLong E.M., DeLong E.M.,  
 RA Durbin K.J., Evansgillista C.C., Fierro C., Fierro C., Fierro C.,  
 RA Foster A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,  
 RA Glodok A., Gong F., Correll J.H., Gu Z., Guan P., Harris M.,  
 RA Horlin D., Houston K.A., Howland T.J., Mel M.-H., Idegami C.,  
 RA Jaitai M., Kalish F., Kaplan G.H., Ke Z., Kenschon J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodish C.D., Kralic C., Kravitz S., Kulp D., Lai Z.,



(2)  
 RN SEQUENCE FROM N.A.  
 RA SU Arabidopsis sequencing project:  
 RA Species: Arabidopsis thaliana EMBL/Genbank/DBJ databases.  
 RA EMBL: AL163553: CAB7278.11.  
 DR InterPro: IPR003103: BAG.  
 DR InterPro: IPR000626: Ubiquitin.  
 DR Pfam: PF02179: BAG\_1.  
 DR SMART: SM00264: BAG\_1.  
 DR PROSITE: PS00053: UBIOUITIN\_2: 1.  
 DR Hypothetical Protein.  
 KW SEQUENCE 303 AA: 3424 MW: EB346BD041A9D37 CRC64;  
 SO  
 Query Match 100.0%; Score 33; DB 10; Length 303;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 14 GCGTSG 19  
 RESULT 11  
 ID 09Y1H0 PRELIMINARY: PRT; 315 AA.  
 AC 09Y1H0;  
 DT 01-NOV-1999 (TRENBLER: 12, Created)  
 DT 01-NOV-1999 (TRENBLER: 12, Last sequence update)  
 DT 01-JUN-2001 (TRENBLER: 17, Last annotation update)  
 OS ARABIDOPSIS THALIANA  
 OS Arabidopsis thaliana  
 OC Eukaryota: Metazoa: Arthropoda: Chelicerata: Arachnida: Acari:  
 OC Parasitiformes: Ixodida: Ixodidae: Amblyomma.  
 OX NCBI\_TaxID:6933;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RA EUNIS J., Hall L., Harbour A.G.;  
 RA Species: Americanus  
 RA Americanus  
 RA Submitted (MAR-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: TO RNA-DIRECTED DNA POLYMERASE (REVERSE  
 CC TRANSCRIPTASE).  
 DR EMBL: AF134586: AAD39461.1: -  
 DR InterPro: IPR004477: RYTS.  
 DR Pfam: PF00078: RYTS.  
 DR SMART: SM00078: RYTS.  
 DR PROSITE: PS00078: RYTS.  
 DR Hypothetical Protein.  
 KW NON TER 315 315  
 FT NON TER 315 315  
 SO SEQUENCE 315 AA: 35459 MW: A10E22ABF3FE14 CRC64;  
 Query Match 100.0%; Score 33; DB 5; Length 315;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 175 GCGTSG 180  
 RESULT 12  
 ID 075953 PRELIMINARY: PRT; 348 AA.  
 AC 075953;  
 DT 01-NOV-1998 (TRENBLER: 08, Created)  
 DT 01-NOV-1998 (TRENBLER: 08, Last sequence update)  
 DT 01-JUN-2001 (TRENBLER: 17, Last annotation update)  
 OS ARABIDOPSIS THALIANA  
 OS Arabidopsis thaliana  
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:  
 OC Euteleostomi: Teleostei: Cyprinodontiformes: Poeciliidae: Gambusia: Gambusia affinis holbrooki

OC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
 OX NCBI\_TaxID:9606;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RA CUSUMANO, A.; LARZIO, A.; ROLL, BOLL, J.L.;  
 RA Species: Homo sapiens  
 RA Homo sapiens  
 RA Submitted (SEP-1998) to the EMBL/Genbank/DBJ databases.  
 CC "differential response of neurons in the hypo family to various  
 CC stress agents in mammalian cells."  
 DR EMBL: AF088982: AAC35860.1: -  
 DR HSSP: P25685: 1HD.  
 DR InterPro: IPR002939: DnaJ\_C.  
 DR InterPro: IPR001633: DnaJ\_N.  
 DR Pfam: PF01556: DnaJ\_C.1.  
 DR SMART: SM00271: DnaJ\_C.1.  
 DR PROSITE: PS00636: DnaJ\_1: 1.  
 DR PROSITE: PS00076: DnaJ\_2: 1.  
 DR Heat shock.  
 KW SEQUENCE 348 AA: 3913 MW: DCGFE45DEPDBCRC CRC64;  
 SO  
 Query Match 100.0%; Score 33; DB 4; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 75 GCGTSG 80  
 RESULT 13  
 ID 09GE43 PRELIMINARY: PRT; 427 AA.  
 AC 09GE43;  
 DT 01-MAR-2001 (TRENBLER: 16, Created)  
 DT 01-MAR-2001 (TRENBLER: 16, Last sequence update)  
 DT 01-JUN-2001 (TRENBLER: 17, Last annotation update)  
 OS ARABIDOPSIS THALIANA  
 OS Arabidopsis thaliana  
 OC Eukaryota: Metazoa: Arthropoda: Chelicerata: Arachnida: Acari:  
 OC Parasitiformes: Ixodida: Ixodidae: Amblyomma.  
 OX NCBI\_TaxID:6933;  
 OX (1)  
 RP SEQUENCE FROM N.A.  
 RA DE LUNA E., Buck W.R., Akiyama H., Arikawa T., Tsubota H.,  
 RA Gonzalez D., Newton A.E., Shaw A.J., Gonzalez D.;  
 RA Species: Arabidopsis thaliana  
 RA Arabidopsis thaliana  
 RA Submitted (MAY-1999) to the EMBL/Genbank/DBJ databases.  
 CC "Phylogenetic analysis of three chloroplast DNA sequences  
 CC from Arabidopsis thaliana and Arabidopsis lyrata."  
 DR EMBL: AF134586: AAD39461.1: -  
 DR InterPro: IPR004477: RYTS.  
 DR Pfam: PF00078: RYTS.  
 DR SMART: SM00078: RYTS.  
 DR PROSITE: PS00078: RYTS.  
 DR Hypothetical Protein.  
 KW NON TER 427 427  
 FT NON TER 427 427  
 SO SEQUENCE 427 AA: 4715 MW: A10E22ABF3FE14 CRC64;  
 Query Match 100.0%; Score 33; DB 5; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 175 GCGTSG 180  
 RESULT 14  
 ID 075953 PRELIMINARY: PRT; 348 AA.  
 AC 075953;  
 DT 01-NOV-1998 (TRENBLER: 08, Created)  
 DT 01-NOV-1998 (TRENBLER: 08, Last sequence update)  
 DT 01-JUN-2001 (TRENBLER: 17, Last annotation update)  
 OS ARABIDOPSIS THALIANA  
 OS Arabidopsis thaliana  
 OC Eukaryota: Metazoa: Chordata: Vertebrata: Euteleostomi:  
 OC Euteleostomi: Teleostei: Cyprinodontiformes: Poeciliidae: Gambusia: Gambusia affinis holbrooki

FT NON\_TER 1 1  
 FT NON\_TER 427 427  
 SQ SEQUENCE 427 AA: 47365 MW: 7F837679470879E CRC64:

Query Match 100.0% Score 33, DB 8; Length 427;  
 Best Local Similarity 100.0%; Pred. No. 2, 9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSC 6  
 DB 383 GCGTSC 388

RESULT 14  
 ID 037M3 PRELIMINARY; PRT: 431 AA.

AC 037M3  
 DT 01-NOV-2000 (TREMBLrel. 13, Created)  
 DT 01-NOV-2000 (TREMBLrel. 13, Last sequence update)  
 DE RUBISCO BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
 LARGE SUBUNIT) (FRAGMENT).  
 GN RECL.  
 OS Solanella denticulata.  
 OC Chloroplast.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Solanaceae; Solanaceae; Solanaceae; Solanaceae;  
 CC NCBI\_TaxID=50508;  
 RN [1]

RA Keroll P., Therrien J.P., Kenrick P.;  
 RT Phylogeny of Selaginellaceae: evaluation of generic/subgeneric  
 relationships based on 18S rDNA sequences.  
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).  
 CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-  
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
 CC -1- ACTIVE SITE (BY SIMILARITY).  
 CC -1- PHOSPHO-D-GLYCERATE. D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-  
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-  
 CC -1- PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.  
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.  
 CC HSP80; P00880.1; HSP80; P00880.1; HSP80; P00880.1;  
 DR InterPro: IPR000685; Rubisco\_Large.  
 DR Pfam: PF00016; Rubisco\_Large.1.  
 DR Prosite: PS00157; Rubisco\_Large.1.  
 KW Carbon dioxide fixation; Chloroplast; lase; Monooxygenase;  
 KW Oxidoreductase; Photorepiration; Photosynthesis.  
 FT NON\_TER 431 431  
 FT NON\_TER 431 431  
 SQ SEQUENCE 431 AA: 46651 MW: EB671D290A96C8E5 CRC64:

Query Match 100.0% Score 33; DB 8; Length 431;  
 Best Local Similarity 100.0%; Pred. No. 2, 9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSC 6  
 DB 382 GCGTSC 387

RESULT 15  
 O03112 PRELIMINARY; PRT: 449 AA.  
 ID O03112

37

AC O03112;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DE RUBISCO BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
 LARGE SUBUNIT) (FRAGMENT).  
 GN RECL.  
 OS Lophocolea heterophylla.  
 OC Chloroplast.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Marchantiophyta;  
 CC Jungermanniales; Geocalycineae; Lophocoleaceae; Lophocolea.  
 CC NCBI\_TaxID=3207;  
 RN [1]  
 RA Lewis L.A., Mishler B.D., Vilgalys R.;  
 RT Phylogenic. Evol. 0:0-0(1997).  
 CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-  
 RUBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
 CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
 THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATION PROCESS. BOTH  
 REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
 CC -1- ACTIVE SITE  
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) = 2 3-  
 CC -1- PHOSPHO-D-GLYCERATE.  
 CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) = 3-  
 CC -1- PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.  
 CC -1- SUBUNIT: 8 LARGE CHAINS + 8 SMALL CHAINS.  
 CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.  
 DR EMBL: U87076; AAC04847.1; -;  
 DR Mendel: 23015; Lopho; rbcL; 23015.  
 DR InterPro: IPR000685; Rubisco\_Large.  
 DR Pfam: PF00016; Rubisco\_Large.1.  
 DR Prosite: PS00157; Rubisco\_Large.1.  
 KW Carbon dioxide fixation; Chloroplast; lase; Monooxygenase;  
 KW Oxidoreductase; Photorepiration; Photosynthesis.  
 FT NON\_TER 1 1  
 FT ACT\_SITE 190 190  
 FT ACT\_SITE 191 191  
 FT NON\_TER 449 449  
 SQ SEQUENCE 449 AA: DE84D95A46D77688 CRC64:

Query Match 100.0% Score 33; DB 8; Length 449;  
 Best Local Similarity 100.0%; Pred. No. 2, 9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSC 6  
 DB 392 GCGTSC 397

RESULT 16  
 ID 031681 PRELIMINARY; PRT: 451 AA.  
 AC 031681  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DE RUBULOSE BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
 LARGE SUBUNIT) (FRAGMENT).  
 GN RECL.  
 OS Androsace rupestris.  
 OC Chloroplast.  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;  
 CC Androsaceae; Androsaceae; Androsaceae; Androsaceae;  
 CC NCBI\_TaxID=13797;  
 RN [1]  
 RA MEDLINE-94356259; PubMed-8075631;  
 RA Mannhart J.R.;  
 RT Phylogenetic analysis of green plant rbcL sequences.\*;  
 RT Mol. Phylogenet. Evol. 3:116-127(1994).

CC -1- FUNCTION: RUBISCO CATALYSES TWO REACTIONS: THE CARBOXYLATION OF D-  
RIBULOSE 1,5-BISPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC  
CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF  
THE PENTOSE SUBSTRATE IN THE PHOTORESPIRATORY PROCESS. BOTH  
REACTIONS OCCUR SIMULTANEOUSLY AND IN COMPETITION AT THE SAME  
CATALYTIC SITE.  
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + CO(2) -> 2 3-  
PHOSPHO-D-GLYCERATE.  
CC -1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BISPHOSPHATE + O(2) -> 3-  
PHOSPHO-D-GLYCERATE + 2-PHOSPHOGLYCOLATE.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- SUBCELLULAR LOCATION: CHLOROPLAST.  
CC -1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.  
DR EMBL: L13473; AA553403.1; -.  
DR Mendel: 3036; Andru: rbel.3036.  
DR Interpro: IPR000685; RUBISCO\_large.  
DR Interpro: IPR001412; RNA-syn.1.  
DR Pfam: PF00016; RUBISCO\_large.1.  
DR Pfam: PF00016; RUBISCO\_large.1.  
DR PROSITE: PS00157; RUBISCO\_LARGE\_1; UNKNOWN\_1.  
KW Carbon dioxide fixation: Chloroplast; Lyase; Monooxygenase;  
KW Oxidoreductase; Photorepiration; Photosynthesis.  
FT NON\_TER 1 1  
FT ACT\_SITE 192 192 BINDING OF CO(2) ACTIVATES THE ENZYME.  
FT ACT\_SITE 193 193 BINDING OF CO(2) ACTIVATES THE ENZYME.  
FT NON\_TER 451 451  
SQ SEQUENCE 451 AA: 45015 MW: 6598713871C7B6C5 CRC64;

Query Match 100.0%; Score 33; DB 8; Length 451;  
Best Local Similarity 100.0%; Pred. No. 2.9e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTSC 6  
DB 394 GCGTSC 399

RESULT 17  
ID 09BCK9 PRELIMINARY; PRT; 468 AA.  
AC 09BCK9; 2001 (TREMURel. 17, Created)  
DT 01-JUN-2001 (TREMURel. 17, Last sequence update)  
DE RIBULOSE-1,5-BISPHOSPHATE CARBOXYLASE/OXYGENASE LARGE SUBUNIT  
(FRAGMENT).  
DR EMBL: A402954; CA004312.1; -.  
DR Interpro: IPR000685; RUBISCO\_large.  
DR Interpro: IPR00157; RUBISCO\_LARGE\_1.  
KW Carbon dioxide fixation: Chloroplast; Lyase; Monooxygenase;  
KW Oxidoreductase; Photorepiration; Photosynthesis.  
FT NON\_TER 1 1  
FT ACT\_SITE 469 469  
SQ SEQUENCE 468 AA: 51314 MW: 614170230F8A2D45 CRC64;

Query Match 100.0%; Score 33; DB 8; Length 468;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGTSC 6  
||||||

DB 386 GCGTSC 391

RESULT 18  
ID 09GHM8 PRELIMINARY; PRT; 469 AA.  
AC 09GHM8; 2001 (TREMURel. 16, Created)  
DT 01-MAR-2001 (TREMURel. 16, Last sequence update)  
DE RIBULOSE 1,5-BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
LARGE SUBUNIT) (FRAGMENT).  
DR EMBL: A402954; CA004312.1; -.  
DR Interpro: IPR000685; RUBISCO\_large.  
DR Interpro: IPR00157; RUBISCO\_LARGE\_1.  
KW Carbon dioxide fixation: Chloroplast; Lyase; Monooxygenase;  
KW Oxidoreductase; Photorepiration; Photosynthesis.  
FT NON\_TER 1 1  
FT ACT\_SITE 469 469  
SQ SEQUENCE 469 AA: 51914 MW: 614693698AD95FE CRC64;

Query Match 100.0%; Score 33; DB 8; Length 469;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTSC 6  
DB 396 GCGTSC 401

RESULT 19  
ID 09HJJO PRELIMINARY; PRT; 470 AA.  
AC 09HJJO; 2001 (TREMURel. 16, Created)  
DT 01-MAR-2001 (TREMURel. 16, Last sequence update)  
DE RIBULOSE 1,5-BISPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO  
LARGE SUBUNIT) (FRAGMENT).  
DR EMBL: A402954; CA004312.1; -.  
DR Interpro: IPR000685; RUBISCO\_large.  
DR Interpro: IPR00157; RUBISCO\_LARGE\_1.  
KW Carbon dioxide fixation: Chloroplast; Lyase; Monooxygenase;  
KW Oxidoreductase; Photorepiration; Photosynthesis.  
FT NON\_TER 1 1  
FT ACT\_SITE 469 469  
SQ SEQUENCE 469 AA: 51914 MW: 614693698AD95FE CRC64;

Query Match 100.0%; Score 33; DB 8; Length 469;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCGTSC 6  
||||||

[illegible]

```

Query Match      100.0%; Score 33; DB 4; Length 470;
Best Local Similarity 100.0%; Pred. No. 3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

Db 26 GGCTSG 31

RESULT	20
ID	09TM74
DB	PRELIMINARY;
DT	01-MAY-2000
DT	01-MAY-2000 (T-PMU13, 13
DT	01-MAY-2000 (TREMUR, 13
DT	01-JUN-2001 (TREMUR, 17, last
DT	01-JUN-2001 (TREMUR, 17, last
DE	RUBULOSE BIPHOSPHATE CARBOXYLASE LARGE CHAIN (EC 4.1.1.39) (RUBISCO
DE	LARGE SUBUNIT).
GN	RBC.
OS	Rhynchosostegium pallidifolium.
OS	Rhynchosostegium pallidifolium.
OC	Embryophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC	Protophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OC	Protophyta; Viridiplantae; Streptophyta; Embryophyta; Bryophyta;
OX	NCBI_Taxid=90343;
OX	NCBI_Taxid=90343;
RP	SEQUENCE FROM N.A.
RP	STRAIN-HIGASHI-HIROSHIMA;
RA	TANAKA H., NAKAO N., ARIKAWA T., YAMAGUCHI T., HIGUCHI M.,
RA	YAMAGUCHI M., SAKAI T.,
RA	YAMAGUCHI M., SAKAI T.,
RL	Chloroplast rbcL sequence data.
RL	Chloroplast rbcL sequence data.
RT	Bryol. Res. 7:233-248(1999).
CC	-1- FUNCTION: RUBISCO CATALYZES TWO REACTIONS: THE CARBOXYLATION OF D-
CC	RUBULOSE 1,5-BIPHOSPHATE, THE PRIMARY EVENT IN PHOTOSYNTHETIC
CC	CARBON DIOXIDE FIXATION, AS WELL AS THE OXIDATIVE FRAGMENTATION OF
CC	THE PEROXIDE SUBSTRATE IN THE PHOTOREGULATION PROCESS. BOTH
CC	REACTIONS ARE SIMULTANEOUSLY AND IN COMPETITION AT THE SAME
CC	ACTIVE SITE.
CC	-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + CO(2) -> 2-3-
CC	PHOSPHO-D-GLYCERATE.
CC	-1- CATALYTIC ACTIVITY: D-RIBULOSE 1,5-BIPHOSPHATE + O(2) -> 3-
CC	PHOSPHO-D-GLYCERATE + 2-PROSOPHOSCOLATE.
CC	-1- SUBUNIT: 8 LARGE CHAINS + 2 SMALL CHAINS (BY SIMILARITY).
CC	-1- SUBCELLULAR LOCATION: CHLOROPLAST (BY SIMILARITY).
CC	-1- SIMILARITY: BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
CC	EMBL: A002644; BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
CC	EMBL: A002644; BELONGS TO THE RUBISCO LARGE CHAIN FAMILY.
CC	Interpro: IPR000685; RUBISCO, large.
CC	Interpro: IPR000685; RUBISCO, large.
CC	PFam: PF00016; RUBISCO, large.
CC	PFam: PF00016; RUBISCO, large.
CC	PROSITE: PS00151; RUBISCO, Chloroplast; 1.
CC	Carbon dioxide fixation; Chloroplast; Lysase; Monooxygenase;
CC	oxidoreductase; Photosynthesis; Photosynthesis; 1.
CC	oxidoreductase; Photosynthesis; Photosynthesis; 1.
CC	SEQUENCE 475 AA: 52740 MW: 407824250kDa157677 CAC64;

```

Query Match          100.0%; Score 33; DB 8; Length 475;
Best Local Similarity 100.0%; Pred. No. 3.1e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0

```

Db 403 GGGTSG 408

RESULT	21	
Q9TM63		
ID	Q9TM63	
AC	Q9TM63.	
	PRELIMINARY;	PRT; 475 AA

Db 403 GGGTSG 408

Accession	Species	Gene	Protein	Length (aa)
U00096.2	<i>Escherichia coli</i>	ompA	OmpA	22
U00096.2	<i>Escherichia coli</i>	ompB	OmpB	22
U00096.2	<i>Escherichia coli</i>	ompC	OmpC	22
U00096.2	<i>Escherichia coli</i>	ompD	OmpD	22
U00096.2	<i>Escherichia coli</i>	ompE	OmpE	22
U00096.2	<i>Escherichia coli</i>	ompF	OmpF	22
U00096.2	<i>Escherichia coli</i>	ompG	OmpG	22
U00096.2	<i>Escherichia coli</i>	ompH	OmpH	22
U00096.2	<i>Escherichia coli</i>	ompI	OmpI	22
U00096.2	<i>Escherichia coli</i>	ompJ	OmpJ	22
U00096.2	<i>Escherichia coli</i>	ompK	OmpK	22
U00096.2	<i>Escherichia coli</i>	ompL	OmpL	22
U00096.2	<i>Escherichia coli</i>	ompM	OmpM	22
U00096.2	<i>Escherichia coli</i>	ompN	OmpN	22
U00096.2	<i>Escherichia coli</i>	ompO	OmpO	22
U00096.2	<i>Escherichia coli</i>	ompP	OmpP	22
U00096.2	<i>Escherichia coli</i>	ompQ	OmpQ	22
U00096.2	<i>Escherichia coli</i>	ompR	OmpR	22
U00096.2	<i>Escherichia coli</i>	ompS	OmpS	22
U00096.2	<i>Escherichia coli</i>	ompT	OmpT	22
U00096.2	<i>Escherichia coli</i>	ompU	OmpU	22
U00096.2	<i>Escherichia coli</i>	ompV	OmpV	22
U00096.2	<i>Escherichia coli</i>	ompW	OmpW	22
U00096.2	<i>Escherichia coli</i>	ompX	OmpX	22
U00096.2	<i>Escherichia coli</i>	ompY	OmpY	22
U00096.2	<i>Escherichia coli</i>	ompZ	OmpZ	22
U00096.2	<i>Escherichia coli</i>	ompA1	OmpA1	22
U00096.2	<i>Escherichia coli</i>	ompA2	OmpA2	22
U00096.2	<i>Escherichia coli</i>	ompA3	OmpA3	22
U00096.2	<i>Escherichia coli</i>	ompA4	OmpA4	22
U00096.2	<i>Escherichia coli</i>	ompA5	OmpA5	22
U00096.2	<i>Escherichia coli</i>	ompA6	OmpA6	22
U00096.2	<i>Escherichia coli</i>	ompA7	OmpA7	22
U00096.2	<i>Escherichia coli</i>	ompA8	OmpA8	22
U00096.2	<i>Escherichia coli</i>	ompA9	OmpA9	22
U00096.2	<i>Escherichia coli</i>	ompA10	OmpA10	22
U00096.2	<i>Escherichia coli</i>	ompA11	OmpA11	22
U00096.2	<i>Escherichia coli</i>	ompA12	OmpA12	22
U00096.2	<i>Escherichia coli</i>	ompA13	OmpA13	22
U00096.2	<i>Escherichia coli</i>	ompA14	OmpA14	22
U00096.2	<i>Escherichia coli</i>	ompA15	OmpA15	22
U00096.2	<i>Escherichia coli</i>	ompA16	OmpA16	22
U00096.2	<i>Escherichia coli</i>	ompA17	OmpA17	22
U00096.2	<i>Escherichia coli</i>	ompA18	OmpA18	22
U00096.2	<i>Escherichia coli</i>	ompA19	OmpA19	22
U00096.2	<i>Escherichia coli</i>	ompA20	OmpA20	22
U00096.2	<i>Escherichia coli</i>	ompA21	OmpA21	22
U00096.2	<i>Escherichia coli</i>	ompA22	OmpA22	22
U00096.2	<i>Escherichia coli</i>	ompA23	OmpA23	22
U00096.2	<i>Escherichia coli</i>	ompA24	OmpA24	22
U00096.2	<i>Escherichia coli</i>	ompA25	OmpA25	22
U00096.2	<i>Escherichia coli</i>	ompA26	OmpA26	22
U00096.2	<i>Escherichia coli</i>	ompA27	OmpA27	22
U00096.2	<i>Escherichia coli</i>	ompA28	OmpA28	22
U00096.2	<i>Escherichia coli</i>	ompA29	OmpA29	22
U00096.2	<i>Escherichia coli</i>	ompA30	OmpA30	22
U00096.2	<i>Escherichia coli</i>	ompA31	OmpA31	22
U00096.2	<i>Escherichia coli</i>	ompA32	OmpA32	22
U00096.2	<i>Escherichia coli</i>	ompA33	OmpA33	22
U00096.2	<i>Escherichia coli</i>	ompA34	OmpA34	22
U00096.2	<i>Escherichia coli</i>	ompA35	OmpA35	22
U00096.2	<i>Escherichia coli</i>	ompA36	OmpA36	22
U00096.2	<i>Escherichia coli</i>	ompA37	OmpA37	22
U00096.2	<i>Escherichia coli</i>	ompA38	OmpA38	22
U00096.2	<i>Escherichia coli</i>	ompA39	OmpA39	22









EMBL: AB015478; BAB11043.1; -  
 InterPro: IP0000172; GMC\_oxred.  
 DR Pfam: PF00732; GMC\_oxred; 1.  
 DR PROSITE: P500624; GMC\_OXRED\_2; 1.  
 KW Lyase.  
 SQ SEQUENCE 563 AA; 61979 MW; 4803807007889P93 CRC64;

Query Match 100.0%; Score 33; DB 10; Length 563;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 52 GCGTSG 57

RESULT 31  
 099Y17 PRELIMINARY; PRT; 565 AA.  
 AC 099Y17; 2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DE PUTATIVE PTS SYSTEM, LACTOSE-SPECIFIC COMPONENT IIIC (SC 2.7.1.69).  
 GN LACE OR SPT1917.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus  
 ON NCBI\_TextID=114;  
 RN [1] TextID=114;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SF370.  
 RX MEDLINE-21192684; PubMed-11295296;  
 RA Perleux J.J., Mesnan W.M., Kilde D.J., Sayle G.J., Lion K.,  
 RA Planteux J.L., Huettenhofer F.Z., Ren O., Zhu H., Song L., White J.,  
 RA Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
 RT \*Complete genome sequence of an M1 strain of Streptococcus pyogenes.\*;  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
 DR EMBL: AE006165; MAK34621.1; -  
 DR Transferrase; Complete proteome.  
 KW SEQUENCE 565 AA; 61209 MW; 122F40043FD2086 CRC64;

Query Match 100.0%; Score 33; DB 2; Length 565;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 471 GCGTSG 476

RESULT 32  
 099S77 PRELIMINARY; PRT; 570 AA.  
 AC 099S77; 2001 (TEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
 DE PTS SYSTEM, LACTOSE-SPECIFIC IIIC COMPONENT.  
 GN LACE OR SA1992.  
 OS Streptococcus aureus subsp. aureus N315  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Streptococcus  
 ON NCBI\_TextID=15879;  
 RN [1] TextID=15879;  
 RP SEQUENCE FROM N.A.  
 RA Kuroda M., Ohta T., Uchiyama T., Baba T., Yuzawa H., Kobayashi T.,  
 RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,  
 RA Matsunaga H., Moriyama A., Murakami H., Hoshiyama A., Mizutani-U T.,  
 RA Takahashi N.K., Sawano T., Inoue K.I., Kato C., Sekimizu K.,

Hirakawa H., Kihara S., Goto S., Yabuzaki J., Kanehisa M.,  
 RA Yamashita A., Oshima K., Furuya K., Yoshino C., Shiba T., Hattori M.,  
 RA Ogawara M., Hayashi H., Hiramatsu K.;  
 RT Whole genome sequencing of methicillin-resistant *Staphylococcus*  
 DR aureus 357-1225-1240(2001).  
 DR EMBL: AP003136; BAB43282.1; -  
 KW Complete proteome.  
 SQ SEQUENCE 570 AA; 62414 MW; 11B6FC1B716028E CRC64;

Query Match 100.0%; Score 33; DB 10; Length 570;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 476 GCGTSG 481

RESULT 33  
 082784 PRELIMINARY; PRT; 574 AA.  
 AC 082784;  
 DT 01-NOV-1998 (TEMBLrel. 08, Created)  
 DT 01-NOV-1998 (TEMBLrel. 08, Last sequence update)  
 DE PUTATIVE PTS SYSTEM, LACTOSE-SPECIFIC COMPONENT IIIC (SC 2.7.1.69).  
 GN LACE OR SPT1917.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus  
 ON NCBI\_TextID=23207;  
 RN [2] TextID=23207;  
 RP SEQUENCE FROM N.A.  
 RA Hu Z., Poulsen J.E.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 RA Hu Z., Poulsen J.E.;  
 RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF043187; AAD0265.1; -  
 DR EMBL: AF043186; AAD0265.1; -  
 DR Mendel: 33350; Pulse:2903/33350  
 DR Mendel: 33350; Pulse:2903/33350  
 DR Mendel: 33350; Pulse:2903/33350  
 DR Mendel: 33350; Pulse:2903/33350  
 DR Pfam: PF00732; GMC\_oxred; 1.  
 DR PROSITE: P500624; GMC\_OXRED\_2; 1.  
 KW Lyase.  
 SQ SEQUENCE 574 AA; 61682 MW; BCDPNEIDCS02539 CRC64;

Query Match 100.0%; Score 33; DB 10; Length 574;  
 Best Local Similarity 100.0%; Pred. No. 3.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 Db 61 GCGTSG 66

RESULT 34  
 P94865 PRELIMINARY; PRT; 577 AA.  
 AC P94865;  
 DT 01-MAY-1997 (TEMBLrel. 03, Created)  
 DT 01-MAY-1997 (TEMBLrel. 03, Last sequence update)  
 DE PUTATIVE PTS SYSTEM, LACTOSE-SPECIFIC COMPONENT IIIC (SC 2.7.1.69).  
 GN LACE.  
 OS Streptococcus pyogenes.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
 OC Streptococcus  
 ON NCBI\_TextID=15879;  
 RN [1] TextID=15879;  
 RP SEQUENCE FROM N.A.  
 RA Kuroda M., Ohta T., Uchiyama T., Baba T., Yuzawa H., Kobayashi T.,  
 RA Cui L., Oguchi A., Aoki K.I., Nagai Y., Lian J., Ito T., Kanamori M.,  
 RA Matsunaga H., Moriyama A., Murakami H., Hoshiyama A., Mizutani-U T.,  
 RA Takahashi N.K., Sawano T., Inoue K.I., Kato C., Sekimizu K.,

GN LACE.  
OS Lactococcus casei.  
OC Bacteria: Firmicutes; Bacillus/Clostridium group; Lactobacillaceae;  
CN AT2C42580; PROTEIN.  
OC Arabidopsis thaliana (Mouse-ear cress).  
OS Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97218691; PubMed-9066115;  
RA Gasolias M., Monedero V., Albert C., Perez-Martinez G.;  
RT Arabidopsis: cloning a model to study the regulation of the lactose operon in  
RT Lactobacillus casei.  
RL FEMS Microbiol. Lett. 148:83-89(1997).  
DR EMBL: Z80834; CAB02556.1; .  
DR InterPro: IPR003352; PTS-ETIC.  
DR InterPro: IPR003501; PTS-ITB.  
DR Pfam: PF03302; PTS-ITB.1.  
DR Pfam: PF03302; PTS-ETIC.1.  
SO SEQUENCE 577 AA; 6269 MW; 32DB7B35E5607CB CRC64;

Query Match 100.0%; Score 33; DB 2; Length 577;  
Best Local Similarity 100.0%; Pred. No. 3; E=0.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Gy 1 GCGTSG 6  
DB 485 GCGTSG 490

RESULT 35  
ID 096V03 PRELIMINARY; PRT; 578 AA.  
AC 096V03;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE ORIGIN RECOGNITION COMPLEX SUBUNIT 2 RELATED PROTEIN.  
OS Neurospora crassa  
OC Neurospora  
OC Eukaryotes: Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
OC Sordariales; Sordariaceae; Neurospora.  
CN MBL\_TaxID-5141;  
OX [1]  
RN SEQUENCE FROM N.A.  
RA Schulte U., Algen V., Hohnsels J., Brandt P., Partmann B., Holland R.,  
RA Neumann G., Meiss H.W., Neumann G.;  
RT Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
RP [12]  
RP German Neurospora genome project;  
RA Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.  
RA EMBL: ALJ53822; CAB88634.1; .  
DR EMBL: ALJ53822; CAB88634.1; .  
DR PROSITE: P50017; G68495.1; UNKNOWN.1  
SO SEQUENCE 578 AA; 63706 MW; E70D0AD01760FC4 CRC64;

Query Match 100.0%; Score 33; DB 3; Length 578;  
Best Local Similarity 100.0%; Pred. No. 3; E=0.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Gy 1 GCGTSG 6  
DB 521 GCGTSG 526

RESULT 36  
ID 09SINI PRELIMINARY; PRT; 618 AA.  
AC 09SINI;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE AT2C42580; PROTEIN.  
CN AT2C42580.  
OC Arabidopsis thaliana (Mouse-ear cress).  
OS Eukaryota: Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20083487; PubMed-10617197;  
RA Lin X., Xian S., Rouniav S.D., Shea T.P., Benito M.-I., Town C.D.,  
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldjanyum T.V.,  
RA Gueli C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,  
RA Adams M.D., Carreira J.J., Sinkov S.E., Mayam L., Tallon L.J., Gill J.E.,  
RA Copenhagen G.P., Preuss D., Nierman W.C., Gooden H.H., Coneryville C.R.,  
RA Salzberg S.L., Fraser C.M., Venter J.C.;  
RT Sequence and analysis of chromosome II of Arabidopsis thaliana.  
RL Nature 402:761-768(1999).  
DR EMBL: AC077681; A022955.1; .  
DR EMBL: AC077681; A022955.1; .  
DR Pfam: PF00155; TPR.6.  
DR SMRT: SM00028; TPR.4.  
SO SEQUENCE 618 AA; 66897 MW; C0E1A18CA212086 CRC64;

Query Match 100.0%; Score 33; DB 10; Length 618;  
Best Local Similarity 100.0%; Pred. No. 4; E=0.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Gy 1 GCGTSG 6  
DB 203 GCGTSG 208

RESULT 37  
ID 056765 PRELIMINARY; PRT; 668 AA.  
AC 056765;  
DT 01-JUN-1998 (TREMblrel. 06, Created)  
DT 01-JUN-1998 (TREMblrel. 06, Last sequence update)  
DT 01-JUN-1998 (TREMblrel. 17, Last annotation update)  
DE ULS2 PROTEIN.  
CN ULS2.  
GN ULS2.  
OS Human cytomegalovirus (strain Towne).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
CN MBL\_TaxID-10363;  
OX [1]  
RN SEQUENCE FROM N.A.  
RA MEDLINE-98241707; PubMed-9573236;  
RA Krosky P.M., Underwood M.R., Turk S.R., Feng K.W.-H., Jain R.K.,  
RA Plak R.G., Westerman A.C., Biron K.K., Townsend L.B., Brach J.C.;  
RT Maps to two open reading frames: UL59 and UL56.\*  
RT \*Resistance of human cytomegalovirus to benzimidazole ribonucleosides  
RT EMBL: AF047521; U142811(1998).  
DR InterPro: IPR002597; Herpesenv.  
DR Pfam: PF01673; Herpesenv.1.  
SO SEQUENCE 668 AA; 74154 MW; 04984BD7A52596CC CRC64;

Query Match 100.0%; Score 33; DB 12; Length 668;  
Best Local Similarity 100.0%; Pred. No. 4; E=0.2;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Gy 1 GCGTSG 6  
DB 422 GCGTSG 427

```

RESULT 38
ID 09ASR6 PRELIMINARY: PRT: 691 AA.
AC 09ASR6: (TREMBL:rel. 17, Created)
DT 01-JUN-2001 (TREMBL:rel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBL:rel. 17, Last annotation update)
DR 01-JUN-2001 (TREMBL:rel. 17, Last annotation update)
DR 01-JUN-2001 (TREMBL:rel. 17, Last annotation update)
OC Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:
OC Eukaryota: Magnoliophyta: eudicotyledons: core eudicotyledons: Rosidae:
OC Eucrosmid II: Brassicales: Brassicaceae: Arabidopsis.
OX NCBI_TaxID:3702;
RN 111
SEQUENCE FROM N.A.
RA Chowik R., Chen H., Kim C.J., Meyers M.C., Shinn P., Banji J.,
RA Bower L., Corns R., Chung M.K., Goldsmith A.D., Hayashizaki Y.,
RA Ishida J., Jones T., Kanya A., Karlin-Neumann G., Kaval J., Lam B.,
RA Lee J.M., Lin J., Liu S.X., Miranda M., Narasaka M., Nguyen M.,
RA Palm C.J., Phan P.K., Quach H.L., Sakano H., Sakurai T., Satou M.,
RA Seki M., Southeick A., Toriumi M., Yamada K., Yu G., Yoshizaki K.,
RA Arabidopsis cDNA clones. "Sexer J.R.":
RA Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
RL EMBL: AF673731; AAK37908.1;
SQ SEQUENCE 691 AA: 75461 MW: A8F8690DB86E19 CRC64;

```

Query Match 100.0%; Score 33; DB 10; Length 691;

Best Local Similarity 100.0%; Pred. No. 47e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GCGTSG 6
Db 203 GCGTSG 208

```

## RESULT 39

```

ID 09DBY4 PRELIMINARY: PRT: 809 AA.

```

```

AC 09DBY4: (TREMBL:rel. 17, Created)
DT 01-JUN-2001 (TREMBL:rel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBL:rel. 17, Last annotation update)
DE 1200009022R1K PROTEIN.
CN 1200009022R1K.

```

OS Mus musculus (Mouse).

OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Eumetazoa:

OC Mammalia: Eutheria: Rodentia: Sciurognathi: Muridae: Murinae: Mus.

OC NCBI\_TaxID:10090;

SEQUENCE FROM N.A.

RA STRAIN=C57BL/6J; TISSUE=LUNG;

RA MEDLINE=1085660; PUBMED=11217851;

RA Aizawa T., Hara A., Fukunishi T., Kono H., Kondo S., Yamada S.,

RA Saito K., Ozaki Y., Gotoh T., Hono H., Kasukawa T., Saito R.,

RA Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Glasl C., King B., Kochava H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schiraldi L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Mashio T.,

RA Sakai K., Ohtsuka T., Furumori M., Aono K., Hatakeyama S.,

RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,

RA Guernier S., Hill D., Hofmann M., Hume D.A., Kanya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombert P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Suzuki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,

RA Watanabe H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilmink L.,

RA Watanabe H., Yoshida K., Hasegawa T., Kawaji H., Kohsaki S.,

RA Watanabe H., Yoshida K., Hasegawa T., Kawaji H., Kohsaki S.,

```

FT "Functional annotation of a full-length mouse cDNA collection."
RL Mature 409:685-690(2001).
DR EMBL: AK004681; BAB23469.1;
DR MGD: MGI:1914123; 1200009022R1K.
DR Interpro: IPR001611; LRR_Cterm.
DR Interpro: IPR000493; LRR_Nterm.
DR Interpro: IPR000493; LRR_Nterm.
DR Interpro: IPR003592; LRR_out.
DR Interpro: IPR003591; LRR_Typ.
DR Pfam: PF00560; LRR_10.
DR Pfam: PF01463; LRRCT_1.
DR PRINTS: PR00019; LRRCTHPT.
DR SMART: SM00020; LRR_8.
DR SMART: SM00022; LRRCT_1.
DR SMART: SM00013; LRRNT_1.
DR SMART: SM00369; LRR_Typ; 11.
SQ SEQUENCE 809 AA: 88809 MW: 5886F55B8A419FA CRC64;

```

Query Match 100.0%; Score 33; DB 11; Length 809;

Best Local Similarity 100.0%; Pred. No. 5.6e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GCGTSG 6
Db 790 GCGTSG 795

```

## RESULT 40

```

ID 09BUL1 PRELIMINARY: PRT: 817 AA.

```

```

AC 09BUL1: (TREMBL:rel. 17, Created)
DT 01-JUN-2001 (TREMBL:rel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBL:rel. 17, Last annotation update)
DE NUCLEAR RECEPTOR NHR-48.
OS Caenorhabditis elegans.

```

OC Eukaryota: Metazoa: Nematoda: Chromadorea: Rhabditida: Rhabditidae.

OC Rhabditidae: Peloderinae: Caenorhabditis.

OC NCBI\_TaxID=6239;

SEQUENCE FROM N.A.

RA Boden A., Maize C.V., Yamamoto K., Cohen F., Sluder A.E.;

RA "Caenorhabditis elegans nuclear receptor sequences exhibit biophysical

RT compatibility with the ligand-binding domain fold."

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL: AF332205; AAK1976.1;

DR Receptor 817 AA: 90706 MW: 036211C9A413AB9 CRC64;

SQ SEQUENCE 817 AA: 90706 MW: 036211C9A413AB9 CRC64;

Query Match 100.0%; Score 33; DB 5; Length 817;

Best Local Similarity 100.0%; Pred. No. 5.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

Qy 1 GCGTSG 6
Db 64 GCGTSG 69

```

Search completed: February 4, 2002, 08:02:21

Job time: 125 sec



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK protein - protein search, using SW model

Run on: February 4, 2002, 08:03:03 ; Search time 38.23 seconds

(without alignments)  
11.625 Million cell updates/sec

Title: us-09-642-660-11

Sequence: 1 GSLGGS 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 522463 seqs, 74073390 residues

Total number of hits satisfying chosen parameters: 522463

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 100%

Listing first 100 summaries

Database:

1: AGeneseq\_1101.\*  
2: /SIDSB/gcgdata/geneseq/geneseq/AA1981.DAT.\*  
3: /SIDSB/gcgdata/geneseq/geneseq/AA1982.DAT.\*  
4: /SIDSB/gcgdata/geneseq/geneseq/AA1983.DAT.\*  
5: /SIDSB/gcgdata/geneseq/geneseq/AA1984.DAT.\*  
6: /SIDSB/gcgdata/geneseq/geneseq/AA1985.DAT.\*  
7: /SIDSB/gcgdata/geneseq/geneseq/AA1986.DAT.\*  
8: /SIDSB/gcgdata/geneseq/geneseq/AA1987.DAT.\*  
9: /SIDSB/gcgdata/geneseq/geneseq/AA1988.DAT.\*  
10: /SIDSB/gcgdata/geneseq/geneseq/AA1989.DAT.\*  
11: /SIDSB/gcgdata/geneseq/geneseq/AA1990.DAT.\*  
12: /SIDSB/gcgdata/geneseq/geneseq/AA1991.DAT.\*  
13: /SIDSB/gcgdata/geneseq/geneseq/AA1992.DAT.\*  
14: /SIDSB/gcgdata/geneseq/geneseq/AA1993.DAT.\*  
15: /SIDSB/gcgdata/geneseq/geneseq/AA1994.DAT.\*  
16: /SIDSB/gcgdata/geneseq/geneseq/AA1995.DAT.\*  
17: /SIDSB/gcgdata/geneseq/geneseq/AA1996.DAT.\*  
18: /SIDSB/gcgdata/geneseq/geneseq/AA1997.DAT.\*  
19: /SIDSB/gcgdata/geneseq/geneseq/AA1998.DAT.\*  
20: /SIDSB/gcgdata/geneseq/geneseq/AA1999.DAT.\*  
21: /SIDSB/gcgdata/geneseq/geneseq/AA2000.DAT.\*  
22: /SIDSB/gcgdata/geneseq/geneseq/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score as high as or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result	Score	Query	Match	Length	ID	Description
1	30	100.0	156	22	AAAM0035	Endoxyloguan tri
2	30	100.0	277	21	AAAY8259	Streptococcus
3	30	100.0	281	21	AAAY8259	Streptococcus
4	30	100.0	287	22	AAAM0097	Endoxyloguan tri
5	30	100.0	411	21	AAAB1097	M. crenulata hemoc
6	30	100.0	411	22	AAAB1099	M. crenulata hemoc
7	30	100.0	416	21	AAAB1002	M. crenulata hemoc
8	30	100.0	416	21	AAAB1002	M. crenulata hemoc
9	30	100.0	416	22	AAAB1157	M. crenulata KLM2
10	30	100.0	433	21	AAAG3306	Zee mays protein f
11	30	100.0	484	21	AAAG3305	Zee mays protein f

12	30	100.0	931	17	AAAM4867	Transferrin bindin
13	30	100.0	1346	18	AAAM2346	Rat tumour suppress
14	30	100.0	1596	18	AAAM2347	Rat tumour suppress
15	28	93.3	21	15	AAAB7055	HTLV peptide. HTLV
16	28	93.3	19	12	AAAG1209	Human secreted pro
17	28	93.3	119	12	AAAG1477	Human colon cancer
18	28	93.3	633	13	AAAG1577	Amino acid sequenc
19	28	93.3	1892	21	AAAG5097	Human secreted pro
20	28	93.3	1892	21	AAAG5097	Human secreted pro
21	28	93.3	1940	21	AAAG5095	Arbidopsis thalia
22	28	93.3	1944	21	AAAG5095	Arbidopsis thalia
23	28	93.3	67	20	AAAY0455	Mycobacterium spec
24	27	90.0	91	21	AAAG2740	Zee mays protein f
25	27	90.0	121	22	AAAB5578	Human secreted pro
26	27	90.0	126	23	AAAB5578	Human secreted pro
27	27	90.0	163	19	AAAM3009	Human protein sequ
28	27	90.0	173	20	AAAY5432	Chlamydia pneumon
29	27	90.0	188	21	AAAG2487	Arbidopsis thalia
30	27	90.0	190	21	AAAB3961	Human secreted pro
31	27	90.0	192	21	AAAG1430	Arbidopsis thalia
32	27	90.0	196	19	AAAB1735	M. tuberculosis im
33	27	90.0	196	19	AAAB1735	Myobacterium tube
34	27	90.0	196	20	AAAY3916	M. tuberculosis im
35	27	90.0	196	20	AAAY3916	M. tuberculosis im
36	27	90.0	233	21	AAAB6951	Human prostate can
37	27	90.0	235	20	AAAY0687	Human prostate can
38	27	90.0	246	19	AAAM3008	Mus musculus I-mfa
39	27	90.0	251	19	AAAM3008	Mus musculus I-mfa
40	27	90.0	280	21	AAAY3925	Consensus eukaryot
41	27	90.0	308	22	AAU12221	Human POU40 poly
42	27	90.0	318	18	AAAM4200	Human mature recep
43	27	90.0	318	18	AAAM4200	Human mature recep
44	27	90.0	318	18	AAAM3754	Human mature recep
45	27	90.0	322	21	AAAG1429	Arbidopsis thalia
46	27	90.0	326	20	AAAB0686	Human secreted pro
47	27	90.0	332	21	AAAY2130	Human mature recep
48	27	90.0	332	21	AAAY2130	Human mature recep
49	27	90.0	336	21	AAAG1428	Arbidopsis thalia
50	27	90.0	336	21	AAAG1428	Arbidopsis thalia
51	27	90.0	337	22	AAAB2935	Human protein sequ
52	27	90.0	340	18	AAAM4199	Human soluble rece
53	27	90.0	340	18	AAAM4199	Human soluble rece
54	27	90.0	349	16	AAAB2582	Human RAGE polyep
55	27	90.0	349	16	AAAB2582	Human RAGE polyep
56	27	90.0	354	20	AAAY3543	Chlamydia thalia
57	27	90.0	354	20	AAAY3543	Chlamydia thalia
58	27	90.0	358	21	AAAB1834	Plasmodium falcip
59	27	90.0	378	21	AAAY5919	Human GAK protein
60	27	90.0	394	22	AAAB8099	S. epidermidis ope
61	27	90.0	404	22	AAAB1925	Extracellular cpr
62	27	90.0	404	22	AAAB1925	Extracellular cpr
63	27	90.0	449	22	AAAB9409	Human secreted pro
64	27	90.0	449	22	AAAB9409	Human secreted pro
65	27	90.0	493	21	AAAG2056	Human ORX ORF1603
66	27	90.0	494	21	AAAG1839	Amino acid sequenc
67	27	90.0	501	22	AAAB6781	Arbidopsis thalia
68	27	90.0	501	22	AAAB6781	Arbidopsis thalia
69	27	90.0	510	22	AAAB5080	Insect steroid rec
70	27	90.0	513	15	AAAB4065	Insect steroid rec
71	27	90.0	513	15	AAAB4065	Insect steroid rec
72	27	90.0	513	18	AAAB18876	D. melanogaster chr
73	27	90.0	513	18	AAAB18876	D. melanogaster chr
74	27	90.0	584	12	AAAB1670	Truncated poly Ig-
75	27	90.0	584	12	AAAB1671	Truncated poly Ig-
76	27	90.0	629	22	AAAB0221	C-terminally trunc
77	27	90.0	629	22	AAAB0221	C-terminally trunc
78	27	90.0	656	22	AAAB9382	Amino acid sequenc
79	27	90.0	657	21	AAAB6407	Human colon cancer
80	27	90.0	673	21	AAAY9544	Human colon cancer
81	27	90.0	708	22	AAAG75616	Human colon cancer
82	27	90.0	773	17	AAAG03177	Rabbit poly-immuno
83	27	90.0	900	20	AAAY3038	A multifunctional
84	27	90.0	900	22	AAAG2036	Saccharomyces cere

ALIGNMENTS

85	27	90.0	900	22	AA820177	S. cerevisiae mult
86	27	90.0	900	22	AA820178	S. cerevisiae mult
87	27	90.0	900	22	AA820179	S. cerevisiae mult
88	27	90.0	967	21	AA845536	Virulence gene pro
89	27	90.0	985	20	AA884519	Moraxella catarrha
90	27	90.0	1005	20	AA884519	Moraxella catarrha
91	27	90.0	1005	20	AA884519	Moraxella catarrha
92	27	90.0	1305	20	AA894418	Moraxella catarrha
93	27	90.0	1305	20	AA894418	Moraxella catarrha
94	27	90.0	1311	20	AA894418	Moraxella catarrha
95	27	90.0	1342	12	AA815833	Protein kinase GAK
96	27	90.0	1342	12	AA815833	Protein kinase GAK
97	27	90.0	1342	12	AA815833	Protein kinase GAK
98	27	90.0	1342	12	AA815833	Protein kinase GAK
99	27	90.0	1342	12	AA815833	Protein kinase GAK
100	27	90.0	1342	12	AA815833	Protein kinase GAK

make it a better or more efficient producer of one or more fine chemicals. Mutagenesis of one or more CMRPs may also result in CMRPs having altered activities which indirectly impact the production of one or more desired fine chemicals from plants. Primers AA88705 - AA88707 are used in the sequencing of the CMR cDNA sequences of the invention.

Sequence 156 AA:

Query Match 100.0%; Score 30; DB 22; Length 156;  
Best Local Similarity 100.0%; Pred. No. 1,3e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSIGGS 6  
111111  
Db 2 gsl9gs 7

## RESULT 2

AA82659 standard; Protein: 277 AA.

AA82659;

10-AUG-2000 (first entry)

Tick salivary gland antigen related immunogen protein sequence.

Tick: vaccine; Infection: salivary gland antigen; Immunogen;

serine protease; cysteine protease; blood sucker.

Haemaphysalis longicornis.

Key Location/Qualifiers

Misc-difference 268 /note= "encoded by T"

JP2000083677-A.

28-MAR-2000.

17-SEP-1998; 98JP-0281932.

17-SEP-1998; 98JP-0281932.

(FARB ) BAYER KK.

WPI: 2000-296340/26.

N-PSDB: AA82654.

A gene encoding tick salivary gland antigen - useful as a vaccine for protecting animals from tick-carried infections

Claim 2; Page 7; 29pp; Japanese.

The present sequence represents a tick salivary gland antigen related immunogen. The present invention also discloses a tick salivary gland antigen encoding any of the above proteins can be used in a vaccine against tick carried infections for domestic animals such as cattle.

Sequence 277 AA;

QY 1 GSIGGS 6

111111

Db 39 gsl9gs 44

Query Match 100.0%; Score 30; DB 21; Length 277;  
Best Local Similarity 100.0%; Pred. No. 2,4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSIGGS 6

111111

Db 39 gsl9gs 44

QY 1 GSIGGS 6

111111

Db 39 gsl9gs 44



```

RESULT 3
AAV90256 standard: Protein: 281 AA.
XX ID AAV90256 standard: Protein: 281 AA.
XX AC AAV90256:
XX 19-SEP-2000 (first entry)
XX Streptococcus equi fibronectin binding protein.
XX Fibronectin binding protein; SFS; vaccine; horse; strangles; therapy;
XX equine upper respiratory tract disease; S. equi infection.
XX Streptococcus equi.
XX Key Location/Qualifiers
FH Misc-difference 23..30
FH /note= "encoded by GTGGCA"
FT Misc-difference 46..54
FT /note= "encoded by TCCGACAG"
FT Misc-difference 70..72
FT /note= "encoded by TCACACGCA AAGAGATTA CTTCAGCCT"
FT /note= "AACCTTCAG AT"
FT Misc-difference 75..83
FT /note= "encoded by ATGCGCCG"
FT Misc-difference 98..106
FT /note= "encoded by CCTGACCA"
FT Misc-difference 122..123
FT /note= "encoded by CAGAGGCTGA GAGAGAGTAA CAGCTCTGGT"
FT /note= "TCAGGTCAAC AC"
FT Misc-difference 127..134
FT /note= "encoded by CAGAT"
FT Misc-difference 150..158
FT /note= "encoded by GCAGAAAA"
FT Misc-difference 174..175
FT /note= "encoded by GGTCAGAGT TGACTGTAA TATCAAGC"
FT /note= "CCTCCTTGC TT"
FT Misc-difference 179..187
FT /note= "encoded by CAGACAGAG"
FT Misc-difference 202..210
FT /note= "encoded by TTAGATCT"
FT Misc-difference 226..227
FT /note= "encoded by GATATGATG ATTCACCAAA ACTTAAAGAC"
FT /note= "GACACCCCT AC"
FT Misc-difference 233..239
FT /note= "encoded by CCTAT"
FT Misc-difference 255..262
FT /note= "encoded by CCAAG"
FT Misc-difference 278..279
FT /note= "encoded by GCGATGAG GTGTGAGAG AGCTACCCG"
FT /note= "INTCTTCA AT"
XX MO200037496-A1.
XX 29-JUN-2000.
XX 21-DEC-1999; 99MO-SE02448.
XX 22-DEC-1998; 98SF-0004491.
XX (GUSG/) GUSS B.
XX (LIND/) LINDMARK H.
XX (JACO/) JACOBSSON K.
XX (FRK/) FRYBERG L.
XX GusB B, Lindmark H, Jacobsson K, Fryberg L;
XX WPI: 2000-442641/38.
XX N-PDB: AAA0869.
XX New protein useful for preparation of vaccines for treatment of
XX strangles caused by Streptococcus equi infection, is able to bind to
XX mammalian fibronectin -

```

```

XX Claim 2: Page 29-30: 34pp: English.
XX This sequence represents the Streptococcus equi fibronectin binding
XX protein (SFS) which binds specifically to
XX fibronectin or its analogues or fragments. The protein, its
XX analogues or fragments may be used for the preparation of a vaccine that
XX protects horses against strangles (a world-wide distributed and serious
XX disease of the equine upper respiratory tract) caused by S. equi
XX infection. The antibody and/or antiserum may also be used for the
XX prophylactic or therapeutic treatment of S. equi infection in animal
XX especially horses. The use of vaccines against S. equi infection
XX binds to fibronectin and provides effective protection against S. equi infections,
XX with fewer side effects.
XX Sequence 281 AA:
XX
XX Query Match 100.0%; Score 30; DB 21; Length 281;
XX Seq. Id. Similarity 100.0%; Posit. No. 24e+02;
XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX Oy 1 GSLGGS 6
XX |||||
XX Db 40 gslggs 45
XX
XX RESULT 4
XX AAM00097 standard: Protein: 287 AA.
XX ID AAM00097
XX AC AAM00097:
XX 28-SEP-2001 (first entry)
XX Endoxylglucan transferase sequence #165.
XX DE
XX Mosa: carbohydrate metabolism related protein; CMRP; sugar; cofactor;
XX fine chemical production; carbohydrate; polysaccharide.
XX Physcomitrella patens.
XX MO200144476-A2.
XX 21-JUN-2001.
XX 14-DEC-2000; 2000MO-EP12697.
XX 16-DEC-1999; 99US-0171101.
XX (BADI ) BAKP PLANT SCI CMH.
XX Lerchl J, Renz A, Einhardt T, Reindl A, Clrpus P, Bischoff F;
XX Frank M, Freund A, Duweng E, Schmidt R, Reski R;
XX WPI: 2001-388155/42.
XX N-PDB: AAB8783.
XX Novel moss nucleic acid molecules encoding a carbohydrate metabolism
XX related protein useful for modulating production of fine chemicals such
XX as carbohydrates, cofactors and enzymes from microorganisms and plants
XX
XX Claim 30: Page 127-128: 133pp: English.
XX This invention relates to nucleic acid molecules AAB88708 - AAB88795
XX isolated from Physcomitrella patens (a moss), which encode carbohydrate
XX metabolism related proteins (CMRP) represented in AAM00022 - AAM00110.
XX Included in the invention is a vector containing the CMRP cDNA, and a
XX host cell transformed with the vector. The host cell (a microorganism,
XX Corynebacterium or Brevibacterium, mosn) as carbohydrate, cofactors
XX useful for production of fine chemicals such as carbohydrates, cofactors
XX and/or enzymes. The nucleic acid molecules are suitable for modifying a

```



Query Match 100.0%; Score 30; DB 22; Length 411;  
 Best Local Similarity 100.0%; Pred. No. 3.6e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 |||||  
 DB 390 gsl9gs 395

## RESULT 7

AA01003 ID AAB1003 standard; Protein: 416 AA.

AC AAB1003;

DT 16-FEB-2001 (first entry)

DE M. crenulata hemocyanin K1H2 domain c'.

KW Hemocyanin; cytosolic; virucide; antibacterial; antiparasitic;

KW Immunomodulatory; antihypertension; antiparasitic; gene therapy;

KW breast; bronchi; colon-rectum; hypertension; vaccine; cocaine misuse;

KW pharmaceutical carrier.

OS Megathura crenulata.

XX WO20005192-A2.

XX 21-SEP-2000.

XX 17-MAR-2000; 2000WO-EP02410.

XX 17-MAR-1999; 99DE-1011971.

XX 20-AUG-1999; 99DE-1039578.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Markl J., Altenhein B., Lieb B., Stiefel T;

XX WPI: 2000-587517/55.

XX New nucleic acid encoding hemocyanin, useful for gene therapy of tumors

XX and for recombinant production of fusion proteins for vaccination

XX Claim 21; Page 142-143; 163pp; German.

XX This invention describes a novel nucleic acid (I) containing a sequence

XX that encodes hemocyanin (II), a domain of (I) or its fragment with the

XX immunological properties of at least one domain of (II). The product of

XX the invention is a nucleic acid comprising nucleic acids of the invention

XX additionally containing antigen-encoding sequences, are useful in gene

XX therapy of tumors. Polypeptides encoded by (I) are useful for treating

XX parasitic or viral infections and tumors, particularly schistosomiasis

XX and carcinoma (of bladder, epithelium, ovary, breast, bronch or

XX colon-rectum), also hypertension, as vaccines, for treating cocaine

XX abuse and very generally as carriers generate antibodies (Ab). Probes

XX based on (I) and Ab are useful for detecting tumor-specific DNA in a cell

XX (by detecting specific binding to cellular DNA or proteins), particularly

XX where associated with the types of carcinoma listed above. Hemocyanins

XX can be produced recombinantly, relatively inexpensively and in adequate

XX amounts, eliminating the need to culture gastropods. When used as a

XX carrier, (II) significantly increases the half-life of the attached

XX pharmaceutical, by inhibiting ultrafiltration in the kidneys.

XX Sequence 416 AA;

XX Query Match 100.0%; Score 30; DB 21; Length 416;

XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 |||||  
 DB 398 gsl9gs 403

## RESULT 8

AA07115 ID AAB7115 standard; Protein: 416 AA.

AC AAB7115;

DT 24-AUG-2001 (first entry)

DE M. crenulata K1H2 domain c'.

KW HH1; K1H1; hemocyanin; cytosolic; vasotropic; vaccine; gene therapy;

KW tumor; parasite infection; viral infection; epithelial carcinoma;

KW abdominal distension; primary carcinoma; bronchial carcinoma;

KW colon carcinoma; cocaine addiction.

OS Megathura crenulata.

XX WO200114536-A2.

XX 01-MAR-2001.

XX 21-AUG-2000; 2000WO-EP08129.

XX 20-AUG-1999; 99DE-1039578.

XX (BIOS-) BIOSYN ARZNEIMITTEL GMBH.

XX Markl J., Altenhein B., Lieb B., Stiefel T;

XX WPI: 2001-191646/19.

XX New nucleic acid encoding hemocyanin, useful for gene therapy of tumors

XX and for recombinant production of fusion proteins for vaccination

XX Claim 25; Page 159-160; 180pp; German.

XX This invention describes a novel nucleic acid molecule (NI) encoding

XX hemocyanin, a hemocyanin domain, or a functional fragment of it with

XX the immunological properties of at least one domain of hemocyanin, and

XX which comprises at least one intron is new. The products of the invention

XX have cytosolic and vasotropic comprising nucleic acids of the invention

XX are useful in treating tumors. Compositions comprising hemocyanin

XX polypeptide are useful in treating parasite and viral infections and as

XX an antitumor agent. It is also useful in treating abnormal blood

XX pressure, bladder carcinomas, epithelial carcinomas, ovarian carcinomas,

XX mammary carcinomas, bronchial carcinomas, and colon carcinomas. It may

XX also be used to cure cocaine addiction.

## RESULT 9

AA07157 ID AAB7157 standard; Protein: 416 AA.

XX Query Match 100.0%; Score 30; DB 22; Length 416;

XX Best Local Similarity 100.0%; Pred. No. 3.7e+02;

XX Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

XX OY 1 GSLOGS 6

XX |||||

XX DB 398 gsl9gs 403

AC AAB71157;  
 XX 24-AUG-2001 (first entry)  
 XX M. crenulata K1H2 domain C.  
 DE  
 XX  
 XX HtH1: K1H1; hemocyanin; cytosolic; vasotrophic; vaccine; gene therapy;  
 XX tumor; parasite infection; viral infection; antitumor agent;  
 XX tumoral blood pressure; bladder carcinoma; epithelial carcinoma;  
 XX ovarian carcinoma; primary carcinoma; bronchial carcinoma;  
 XX colon carcinoma; cocaine addiction.  
 OS  
 XX Megathura crenulata.  
 XX MO200114536-A2.  
 XX  
 XX 01-MAR-2001.  
 PD  
 XX  
 XX 21-AUG-2000: 2000MO-EP08129.  
 PF  
 XX 20-AUG-1999: 99DE-1035578.  
 PR  
 XX (B10S)- B10STR ARZEMWITTEL GMBH.  
 XX  
 XX Markl J, Altenheim B, Lieb B, Stiefel T;  
 PI WPI: 2001-191646/19.  
 DR  
 XX New nucleic acid encoding hemocyanin, useful for gene therapy of tumors  
 PF and for recombinant production of fusion proteins for vaccination -  
 P  
 XX Disclosure: Figure 11: 180pp; German.  
 PS  
 XX This invention describes a novel nucleic acid molecule (NI) encoding  
 CC hemocyanin, a hemocyanin domain, or a functional fragment of it with  
 CC the immunological properties of at least one domain of hemocyanin, and  
 CC the use of at least one fusion is new. The products of the invention  
 CC have cytotoxic and antitumor activity and can be used in vaccines or  
 CC for gene therapy. Compositions comprising at least one of the invention  
 CC are useful in treating tumors. Compositions comprising at least one of the invention  
 CC polypeptide are useful in treating parasite and viral infections and as  
 CC an antitumor agent. It is also useful in treating abnormal blood and  
 CC pressure, bladder carcinoma, epithelial carcinoma, ovarian carcinoma,  
 CC colon carcinoma, bronchial carcinoma, and colon carcinomas. It may  
 CC also be used to cure cocaine addiction.  
 XX  
 XX Sequence 416 AA:  
 SQ  
 QUERY MATCH 100.0%; Score 30; DB 23; Length 416;  
 Bseq Local 100.0%; Pred. No. 3,7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSGCS 6  
 111111  
 Db 398 gslsgs 403  
 RESULT 10  
 ID AAG33306 standard: Protein: 433 AA.  
 AAG33306;  
 AC AAG33306;  
 XX  
 XX 18-OCT-2000 (first entry)  
 XX  
 XX Zea mays protein fragment SEQ ID NO: 40336.  
 DE  
 XX  
 XX Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridization assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 XX Zea mays subsp. mays.

XX  
 FN EPI03405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000: 2000EP-0301439.  
 PF  
 XX 25-FEB-1999: 99US-0121825.  
 PR 25-FEB-1999: 99US-0121825.  
 PR 05-MAR-1999: 99US-0123548.  
 PR 23-MAR-1999: 99US-0125788.  
 PR 25-MAR-1999: 99US-0126264.  
 PR 25-MAR-1999: 99US-0126785.  
 PR 12-MAR-1999: 99US-0127452.  
 PR 08-MAR-1999: 99US-0128214.  
 PR 16-APR-1999: 99US-0129845.  
 PR 19-APR-1999: 99US-0130077.  
 PR 21-APR-1999: 99US-0130449.  
 PR 23-APR-1999: 99US-0130510.  
 PR 28-APR-1999: 99US-0130891.  
 PR 30-APR-1999: 99US-0131449.  
 PR 30-APR-1999: 99US-0132407.  
 PR 04-MAY-1999: 99US-0132484.  
 PR 05-MAY-1999: 99US-0132485.  
 PR 06-MAY-1999: 99US-0132486.  
 PR 07-MAY-1999: 99US-0132487.  
 PR 11-MAY-1999: 99US-0132493.  
 PR 14-MAY-1999: 99US-0134218.  
 PR 14-MAY-1999: 99US-0134219.  
 PR 14-MAY-1999: 99US-0134221.  
 PR 18-MAY-1999: 99US-0134370.  
 PR 19-MAY-1999: 99US-0134368.  
 PR 20-MAY-1999: 99US-0135124.  
 PR 21-MAY-1999: 99US-0135523.  
 PR 24-MAY-1999: 99US-0135629.  
 PR 25-MAY-1999: 99US-0136021.  
 PR 26-MAY-1999: 99US-0136392.  
 PR 01-JUN-1999: 99US-0137052.  
 PR 03-JUN-1999: 99US-0137526.  
 PR 04-JUN-1999: 99US-0137602.  
 PR 07-JUN-1999: 99US-0137724.  
 PR 08-JUN-1999: 99US-0138094.  
 PR 10-JUN-1999: 99US-0138540.  
 PR 11-JUN-1999: 99US-0138644.  
 PR 14-JUN-1999: 99US-0139452.  
 PR 16-JUN-1999: 99US-0139453.  
 PR 17-JUN-1999: 99US-0139452.  
 PR 18-JUN-1999: 99US-0139454.  
 PR 18-JUN-1999: 99US-0139455.  
 PR 18-JUN-1999: 99US-0139459.  
 PR 18-JUN-1999: 99US-0139458.  
 PR 18-JUN-1999: 99US-0139459.  
 PR 18-JUN-1999: 99US-0139460.  
 PR 18-JUN-1999: 99US-0139461.  
 PR 18-JUN-1999: 99US-0139462.  
 PR 18-JUN-1999: 99US-0139463.  
 PR 18-JUN-1999: 99US-0139464.  
 PR 18-JUN-1999: 99US-0139465.  
 PR 21-JUN-1999: 99US-0139817.  
 PR 22-JUN-1999: 99US-0139899.  
 PR 23-JUN-1999: 99US-0140353.  
 PR 23-JUN-1999: 99US-0140354.  
 PR 26-JUN-1999: 99US-0140695.  
 PR 29-JUN-1999: 99US-0140823.  
 PR 30-JUN-1999: 99US-0141287.

PR 01-JUL-1999: 9905-0141842.  
 PR 01-JUL-1999: 9905-0142154.  
 PR 02-JUL-1999: 9905-0142055.  
 PR 06-JUL-1999: 9905-0142390.  
 PR 08-JUL-1999: 9905-0142803.  
 PR 09-JUL-1999: 9905-0142920.  
 PR 13-JUL-1999: 9905-0143542.  
 PR 14-JUL-1999: 9905-0143624.  
 PR 15-JUL-1999: 9905-0144005.  
 PR 16-JUL-1999: 9905-0144085.  
 PR 16-JUL-1999: 9905-0144086.  
 PR 19-JUL-1999: 9905-0144325.  
 PR 19-JUL-1999: 9905-0144332.  
 PR 19-JUL-1999: 9905-0144333.  
 PR 19-JUL-1999: 9905-0144334.  
 PR 19-JUL-1999: 9905-0144335.  
 PR 20-JUL-1999: 9905-0144352.  
 PR 20-JUL-1999: 9905-0144632.  
 PR 20-JUL-1999: 9905-0144884.  
 PR 21-JUL-1999: 9905-0144886.  
 PR 21-JUL-1999: 9905-0145086.  
 PR 21-JUL-1999: 9905-0145088.  
 PR 22-JUL-1999: 9905-0145085.  
 PR 22-JUL-1999: 9905-0145087.  
 PR 22-JUL-1999: 9905-0145089.  
 PR 23-JUL-1999: 9905-0145132.  
 PR 23-JUL-1999: 9905-0145213.  
 PR 23-JUL-1999: 9905-0145214.  
 PR 26-JUL-1999: 9905-0145276.  
 PR 27-JUL-1999: 9905-0145913.  
 PR 27-JUL-1999: 9905-0145918.  
 PR 28-JUL-1999: 9905-0145919.  
 PR 28-JUL-1999: 9905-0145951.  
 PR 28-JUL-1999: 9905-0146381.  
 PR 02-AUG-1999: 9905-0146388.  
 PR 02-AUG-1999: 9905-0146389.  
 PR 04-AUG-1999: 9905-0147038.  
 PR 04-AUG-1999: 9905-0147204.  
 PR 04-AUG-1999: 9905-0147302.  
 PR 05-AUG-1999: 9905-0147192.  
 PR 06-AUG-1999: 9905-0147203.  
 PR 06-AUG-1999: 9905-0147203.  
 PR 06-AUG-1999: 9905-014716.  
 PR 09-AUG-1999: 9905-0147493.  
 PR 09-AUG-1999: 9905-0147935.  
 PR 10-AUG-1999: 9905-0148171.  
 PR 11-AUG-1999: 9905-0148319.  
 PR 12-AUG-1999: 9905-0148361.  
 PR 13-AUG-1999: 9905-0148364.  
 PR 16-AUG-1999: 9905-0148684.  
 PR 17-AUG-1999: 9905-0149368.  
 PR 17-AUG-1999: 9905-0149175.  
 PR 18-AUG-1999: 9905-0149126.  
 PR 20-AUG-1999: 9905-0149722.  
 PR 20-AUG-1999: 9905-0149723.  
 PR 23-AUG-1999: 9905-0149902.  
 PR 23-AUG-1999: 9905-0149930.  
 PR 25-AUG-1999: 9905-0150566.  
 PR 26-AUG-1999: 9905-0150684.  
 PR 27-AUG-1999: 9905-0151065.  
 PR 27-AUG-1999: 9905-0151060.  
 PR 30-AUG-1999: 9905-0151303.  
 PR 31-AUG-1999: 9905-0151438.  
 PR 01-SEP-1999: 9905-0151930.  
 PR 07-SEP-1999: 9905-0152163.  
 PR 10-SEP-1999: 9905-0153070.  
 PR 13-SEP-1999: 9905-0153158.  
 PR 13-SEP-1999: 9905-0154018.

PR 16-SEP-1999: 9905-0154039.  
 PR 20-SEP-1999: 9905-0154779.  
 PR 22-SEP-1999: 9905-0155139.  
 PR 23-SEP-1999: 9905-0155456.  
 PR 24-SEP-1999: 9905-0155456.  
 PR 28-SEP-1999: 9905-0156458.  
 PR 29-SEP-1999: 9905-0156596.  
 PR 04-OCT-1999: 9905-0157117.  
 PR 05-OCT-1999: 9905-0157793.  
 PR 06-OCT-1999: 9905-0157865.  
 PR 07-OCT-1999: 9905-0158029.  
 PR 10-OCT-1999: 9905-0158359.  
 PR 10-OCT-1999: 9905-0158637.  
 PR 13-OCT-1999: 9905-0159293.  
 PR 13-OCT-1999: 9905-0159294.  
 PR 13-OCT-1999: 9905-0159295.  
 PR 14-OCT-1999: 9905-0159339.  
 PR 14-OCT-1999: 9905-0159330.  
 PR 14-OCT-1999: 9905-0159637.  
 PR 14-OCT-1999: 9905-0159638.  
 PR 18-OCT-1999: 9905-0159584.  
 PR 21-OCT-1999: 9905-0160711.  
 PR 21-OCT-1999: 9905-0160767.  
 PR 21-OCT-1999: 9905-0160778.  
 PR 21-OCT-1999: 9905-0160770.  
 PR 21-OCT-1999: 9905-0160815.  
 PR 22-OCT-1999: 9905-0160881.  
 PR 22-OCT-1999: 9905-0160981.  
 PR 22-OCT-1999: 9905-0160989.  
 PR 25-OCT-1999: 9905-0161404.  
 PR 25-OCT-1999: 9905-0161405.  
 PR 25-OCT-1999: 9905-0161406.  
 PR 26-OCT-1999: 9905-0161360.  
 PR 26-OCT-1999: 9905-0161361.  
 PR 28-OCT-1999: 9905-0161920.  
 PR 28-OCT-1999: 9905-0161992.  
 PR 28-OCT-1999: 9905-0161993.  
 PR 29-OCT-1999: 9905-0162142.

Query Match 100.0% Score 30: DB 21: Length 433:  
 Best Local Similarity 100.0% Pred. No. 3. Be:02:  
 Matches 6: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 GSLGS 6  
 DB 2 gslgs 7

RESULT 11  
 AAC33305  
 ID AAC33305 standard: Protein; 484 AA.  
 AC AAC33305:  
 DT 18-OCT-2000 (first entry)  
 XX  
 DE Zea mays protein fragment SD0 ID NO: 40335.  
 XX  
 KW Protein identification: signal transduction pathway; metabolic pathway;  
 KW hyrdatisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence; corn.  
 XX  
 OS Zea mays subsp. mays.  
 PN EP1033405-A2.  
 XX  
 PD 06-SEP-2000.  
 XX  
 PR 25-FEB-2000; 2000EP-0301439.  
 XX

PR 25-FEB-1999; 99US-0121825.  
PR 05-MAR-1999; 99US-0123160.  
PR 23-MAR-1999; 99US-0123548.  
PR 23-MAR-1999; 99US-0123548.  
PR 25-MAR-1999; 99US-0123548.  
PR 29-MAR-1999; 99US-0126785.  
PR 01-APR-1999; 99US-0127462.  
PR 06-APR-1999; 99US-0128234.  
PR 08-APR-1999; 99US-0128714.  
PR 16-APR-1999; 99US-0129845.  
PR 16-APR-1999; 99US-0130042.  
PR 21-APR-1999; 99US-0130510.  
PR 23-APR-1999; 99US-0130891.  
PR 23-APR-1999; 99US-0131449.  
PR 28-APR-1999; 99US-0131449.  
PR 30-APR-1999; 99US-0131407.  
PR 01-MAY-1999; 99US-0132484.  
PR 05-MAY-1999; 99US-0132486.  
PR 06-MAY-1999; 99US-0132487.  
PR 06-MAY-1999; 99US-0132863.  
PR 07-MAY-1999; 99US-0132956.  
PR 11-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134218.  
PR 14-MAY-1999; 99US-0134219.  
PR 14-MAY-1999; 99US-0134370.  
PR 16-MAY-1999; 99US-0134768.  
PR 19-MAY-1999; 99US-0134941.  
PR 19-MAY-1999; 99US-0135124.  
PR 20-MAY-1999; 99US-0135353.  
PR 24-MAY-1999; 99US-0135353.  
PR 25-MAY-1999; 99US-0136029.  
PR 27-MAY-1999; 99US-0136392.  
PR 28-MAY-1999; 99US-0136782.  
PR 01-JUN-1999; 99US-0137222.  
PR 03-JUN-1999; 99US-0137528.  
PR 07-JUN-1999; 99US-0137502.  
PR 08-JUN-1999; 99US-0137674.  
PR 10-JUN-1999; 99US-0138540.  
PR 10-JUN-1999; 99US-0138847.  
PR 14-JUN-1999; 99US-0139119.  
PR 16-JUN-1999; 99US-0139452.  
PR 16-JUN-1999; 99US-0139453.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139454.  
PR 18-JUN-1999; 99US-0139456.  
PR 18-JUN-1999; 99US-0139457.  
PR 18-JUN-1999; 99US-0139458.  
PR 18-JUN-1999; 99US-0139459.  
PR 18-JUN-1999; 99US-0139460.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139461.  
PR 18-JUN-1999; 99US-0139750.  
PR 18-JUN-1999; 99US-0139750.  
PR 21-JUN-1999; 99US-0139817.  
PR 21-JUN-1999; 99US-0139819.  
PR 21-JUN-1999; 99US-0139819.  
PR 23-JUN-1999; 99US-0140253.  
PR 24-JUN-1999; 99US-0140584.  
PR 28-JUN-1999; 99US-0140823.  
PR 29-JUN-1999; 99US-0140991.  
PR 30-JUN-1999; 99US-0141827.  
PR 30-JUN-1999; 99US-0141842.  
PR 01-JUL-1999; 99US-0142154.  
PR 02-JUL-1999; 99US-0142955.  
PR 06-JUL-1999; 99US-0142803.  
PR 08-JUL-1999; 99US-0142920.  
PR 09-JUL-1999; 99US-0142920.  
PR 12-JUL-1999; 99US-0142977.  
  
PR 13-JUL-1999; 99US-0143542.  
PR 14-JUL-1999; 99US-0143542.  
PR 15-JUL-1999; 99US-0144005.  
PR 16-JUL-1999; 99US-0144085.  
PR 16-JUL-1999; 99US-0144086.  
PR 19-JUL-1999; 99US-0144335.  
PR 19-JUL-1999; 99US-0144331.  
PR 19-JUL-1999; 99US-0144332.  
PR 13-JUL-1999; 99US-0144334.  
PR 19-JUL-1999; 99US-0144335.  
PR 20-JUL-1999; 99US-0144332.  
PR 20-JUL-1999; 99US-0144632.  
PR 20-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0144884.  
PR 21-JUL-1999; 99US-0145088.  
PR 22-JUL-1999; 99US-0145085.  
PR 22-JUL-1999; 99US-0145087.  
PR 22-JUL-1999; 99US-0145089.  
PR 22-JUL-1999; 99US-0145192.  
PR 23-JUL-1999; 99US-0145195.  
PR 23-JUL-1999; 99US-0145224.  
PR 26-JUL-1999; 99US-0145226.  
PR 27-JUL-1999; 99US-0145913.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145918.  
PR 27-JUL-1999; 99US-0145919.  
PR 27-JUL-1999; 99US-0145921.  
PR 02-AUG-1999; 99US-0146388.  
PR 02-AUG-1999; 99US-0146388.  
PR 03-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147038.  
PR 04-AUG-1999; 99US-0147302.  
PR 04-AUG-1999; 99US-0147302.  
PR 05-AUG-1999; 99US-0147392.  
PR 06-AUG-1999; 99US-0147393.  
PR 06-AUG-1999; 99US-0147416.  
PR 09-AUG-1999; 99US-0147493.  
PR 09-AUG-1999; 99US-0147935.  
PR 10-AUG-1999; 99US-0148171.  
PR 11-AUG-1999; 99US-0148319.  
PR 11-AUG-1999; 99US-0148361.  
PR 13-AUG-1999; 99US-0148361.  
PR 13-AUG-1999; 99US-0148684.  
PR 16-AUG-1999; 99US-0149366.  
PR 17-AUG-1999; 99US-0149175.  
PR 18-AUG-1999; 99US-0149426.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149722.  
PR 20-AUG-1999; 99US-0149722.  
PR 23-AUG-1999; 99US-0149902.  
PR 23-AUG-1999; 99US-0149930.  
PR 25-AUG-1999; 99US-0150566.  
PR 26-AUG-1999; 99US-0150884.  
PR 27-AUG-1999; 99US-0151065.  
PR 27-AUG-1999; 99US-0151066.  
PR 27-AUG-1999; 99US-0151066.  
PR 30-AUG-1999; 99US-0151103.  
PR 31-AUG-1999; 99US-0151130.  
PR 01-SEP-1999; 99US-0151330.  
PR 07-SEP-1999; 99US-0152363.  
PR 10-SEP-1999; 99US-0153070.  
PR 13-SEP-1999; 99US-0153758.  
PR 13-SEP-1999; 99US-0154018.  
PR 16-SEP-1999; 99US-0154723.  
PR 20-SEP-1999; 99US-0155139.  
PR 22-SEP-1999; 99US-0155139.  
PR 23-SEP-1999; 99US-0155486.  
PR 24-SEP-1999; 99US-0155655.  
PR 28-SEP-1999; 99US-0156456.  
PR 29-SEP-1999; 99US-0156596.

PR 04-OCT-1999; 9905-0157117.  
 PR 05-OCT-1999; 9905-0157121.  
 PR 06-OCT-1999; 9905-0157665.  
 PR 07-OCT-1999; 9905-0158029.  
 PR 08-OCT-1999; 9905-0158232.  
 PR 12-OCT-1999; 9905-0158369.  
 PR 13-OCT-1999; 9905-0159293.  
 PR 13-OCT-1999; 9905-0159294.  
 PR 13-OCT-1999; 9905-0159329.  
 PR 13-OCT-1999; 9905-0159330.  
 PR 14-OCT-1999; 9905-0159331.  
 PR 14-OCT-1999; 9905-0159337.  
 PR 14-OCT-1999; 9905-0159638.  
 PR 18-OCT-1999; 9905-0159584.  
 PR 21-OCT-1999; 9905-0160767.  
 PR 21-OCT-1999; 9905-0160768.  
 PR 21-OCT-1999; 9905-0160770.  
 PR 21-OCT-1999; 9905-0160814.  
 PR 21-OCT-1999; 9905-0160815.  
 PR 22-OCT-1999; 9905-0160980.  
 PR 22-OCT-1999; 9905-0160981.  
 PR 22-OCT-1999; 9905-0161404.  
 PR 25-OCT-1999; 9905-0161405.  
 PR 25-OCT-1999; 9905-0161406.  
 PR 26-OCT-1999; 9905-0161359.  
 PR 26-OCT-1999; 9905-0161360.  
 PR 26-OCT-1999; 9905-0161361.  
 PR 28-OCT-1999; 9905-0161982.  
 PR 28-OCT-1999; 9905-0161993.  
 PR 28-OCT-1999; 9905-0161993.  
 PR 29-OCT-1999; 9905-0162142.

Query Match Similarity 100.0%; Score 30; DB 21; Length 484;  
 Best Local Similarity 100.0%; Pred. No. g: 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSIGS 5  
 DB 53 gsl99s 58

RESULT 12

ID ANM04867 standard; Protein: 931 AA.

XX ANM04867;

AC 19-JAN-1997 (first entry)

DT Transferrin binding protein 1 gene.

KW Transferrin binding protein 1; Tbp1; vaccine; antibody; diagnosis;

KW swine pleuropneumonia.

XX Actinobacillus pleuropneumoniae strain 1371 serotype 1.

OS Actinobacillus pleuropneumoniae

XX Key location/Qualifiers

FT Peptide 1..22 /label= sig-peptide

PN EP73708-A2.

PD 25-SEP-1996.

PF 21-MAR-1996; 96EP-0870033.

PR 24-MAR-1995; 95ES-0000592.

XX (HPR-) LAB HIPRA SA.

XX

PI Daban M, Espuna E, Medrano A, Querol E;

XX WPI; 1996-427056/43.

DR N-PDB; AAT38071.

XX Actinobacillus pleuropneumoniae transferrin binding protein 1 - for

PT detection of antibodies useful diagnostically and in universal

XX vaccine against porcine pleuropneumonia

XX Claim 12; Page 10-17, 22pp; English.

CC Transferrin binding protein 1 (Tbp1) (ANM04867) of Actinobacillus

CC pleuropneumoniae (APP) strain 1370 derived from strain Hpn-1

CC that serves as a channel for transport of iron across the outer

CC membrane. Its amino acid sequence was deduced from a gene

CC (AAT38071) obt'd. from APP genomic DNA. Recombinant Tbp1, or its

CC antigenic fragments, can be produced in transformed host cells,

CC It is used to formulate vaccines against porcine pleuropneumonia

CC to prepare antibodies (useful for serotherapy) and to prepare

CC diagnostic reagents.

XX Sequence 931 AA:

Query Match Similarity 100.0%; Score 30; DB 17; Length 931;

Best Local Similarity 100.0%; Pred. No. g: 0; Mismatches 0; Indels 0; Gaps 0;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSIGS 6

DB 151 gsl99s 156

RESULT 13

ID ANM1346 standard; Protein: 1346 AA.

XX ANM1346;

AC 20-JUL-1998 (first entry)

DT Rat tumour suppressor protein SSeCKs (active truncated form).

KW SSeCKs; tumour suppressor gene; rat; protein kinase C; mitosis;

KW cancer; malignancy; cell proliferation; Alzheimer's disease;

KW therapy.

XX Rattus sp.

OS Rattus sp.

XX Key location/Qualifiers

FT MISC-difference 72 /note= "encoded by TGT"

FT MISC-difference 496 /note= "encoded by GTC"

FT MISC-difference /note= "encoded by GTC"

FT MISC-difference 757 /note= "encoded by CCC"

FT MISC-difference 785 /note= "encoded by AGG"

FT Region 24..32 /note= "glycine-rich region"

FT Peptide 131..134 /note= "nuclear localisation signal"

FT Region 138..141 /note= "glycine-rich region"

FT Peptide 202..205 /note= "nuclear localisation signal"

FT Peptide 289..291 /note= "nuclear localisation signal"

FT Region 306..315 /note= "glycine-rich region"

FT Peptide 368..371  
 FT /note= "nuclear localisation signal"  
 FT Region 103..105  
 FT /note= "zinc finger"  
 PN MO9740059-A1.  
 PD 30-OCT-1997.  
 PR 18-APR-1997: 97MO-US06830.  
 PR 18-JUN-1996: 96US-065401.  
 PR 19-APR-1996: 96US-0635121.  
 XX  
 XX (GELM/) GELMAN I.  
 XX (JAKE/) JAKEN S.  
 XX Gelman I, Jaken S;  
 PI WPI: 1997-535770/49.  
 DR N-PSDB: AAV02301.  
 XX Tumour suppressor gene SSeCKs - used as a mitotic regulator, and  
 XX inhibitor of malignant phenotype  
 XX Example 6: Flg 3A-C; 162pp; English.  
 XX This polypeptide comprises an active truncated form of the novel  
 CC tumour suppressor protein SSeCKs that is a substrate of protein  
 CC kinase and which acts as a negative regulatory of mitosis and as  
 CC an inhibitor of the transformed phenotype. Its amino acid sequence  
 CC was deduced from the complementary DNA sequence (see AAV02301) and  
 CC NIH373 cells. The full-length SSeCKs (see AAV02301) is also  
 CC homologous and hybridising nucleic acids are claimed, as are  
 CC isolated proteins encoded by such nucleic acids, vectors comprising  
 CC the nucleic acids, transformed phenotype in a host cell by introducing  
 CC the nucleic acids, and/or with the expression of a malignant  
 CC product into a host cell inhibits the growth of the host cell,  
 CC allowing the treatment of diseases associated with disorders of  
 CC proliferation and/or with the expression of a malignant phenotype.  
 CC SSeCKs can also be used to treat or identify disorders of  
 CC cytoskeletal structure and cellular architecture (such as  
 CC Alzheimer's disease), and may be a marker for aberrancies in  
 CC fertility and/or nervous system development.  
 XX Sequence 1346 AA:  
 SQ  
 Query Match 100.0%; Score 30; DB 18; Length 1346;  
 Host Local similarity 100.0%; Pred. No. 1.3e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CSLGGS 6  
 Db 960 981996 965  
 |||||  
 RESULT 14  
 ID AAV01347 standard; Protein: 1596 AA.  
 XX AAV01347;  
 XX 20-JUL-1998 (first entry)  
 DE Rat tumour suppressor protein SSeCKs.  
 XX SSeCKs: tumour suppressor gene; rat; protein kinase C; mitosis;  
 XX cancer; malignancy; cell proliferation; Alzheimer's disease;  
 XX therapy.

OS Rattus sp.  
 FT Key  
 FT Peptide 1..10  
 FT Location/Qualifiers  
 FT /note= "myristylation site"  
 FT Region 410..418  
 FT /note= "glycine-rich region"  
 FT Peptide 517..520  
 FT /note= "nuclear localisation signal"  
 FT Region 520..527  
 FT /note= "glycine-rich region"  
 FT Peptide 677..680  
 FT /note= "nuclear localisation signal"  
 FT Peptide 757..760  
 FT /note= "nuclear localisation signal"  
 FT Region 103..105  
 FT /note= "zinc finger"  
 FT Modified-site 279..307  
 FT /note= "proposed protein kinase C phosphorylation site"  
 FT Modified-site 504..526  
 FT /note= "proposed protein kinase C phosphorylation site"  
 FT Modified-site 592..614  
 FT /note= "proposed protein kinase C phosphorylation site"  
 FT Modified-site 741..766  
 FT /note= "proposed protein kinase C phosphorylation site"  
 PN MO9740059-A1.  
 PD 30-OCT-1997.  
 PR 18-APR-1997: 97MO-US06830.  
 PR 18-JUN-1996: 96US-065401.  
 PR 19-APR-1996: 96US-0635121.  
 XX  
 XX (GELM/) GELMAN I.  
 XX (JAKE/) JAKEN S.  
 XX Gelman I, Jaken S;  
 XX WPI: 1997-535770/49.  
 DR N-PSDB: AAV02302.  
 XX Tumour suppressor gene SSeCKs - used as a mitotic regulator, and  
 XX inhibitor of malignant phenotype  
 XX Claim 4: Flg 11A-L; 162pp; English.  
 XX This polypeptide comprises the novel tumour suppressor protein  
 CC SSeCKs that is a substrate of protein kinase C and which acts as a  
 CC negative regulatory of mitosis and as an inhibitor of the  
 CC transformed phenotype. Its amino acid sequence was deduced from an  
 CC complementary DNA sequence (see AAV02302) and is also  
 CC homologous and hybridising nucleic acids are claimed, as are  
 CC isolated proteins encoded by such nucleic acids, vectors comprising  
 CC the nucleic acids, transformed phenotype in a host cell by introducing  
 CC the nucleic acids, and/or with the expression of a malignant  
 CC product into a host cell inhibits the growth of the host cell,  
 CC allowing the treatment of diseases associated with disorders of  
 CC proliferation and/or with the expression of a malignant  
 CC phenotype. SSeCKs can also be used to treat or identify disorders  
 CC of cytoskeletal structure and cellular architecture (such as  
 CC Alzheimer's disease), and may be a marker for aberrancies in  
 CC fertility and/or nervous system development.  
 XX Sequence 1596 AA:  
 SQ



Query Match 100.0%; Score 30; DB 18; Length 1596;  
 Best Local Similarity 100.0%; Pred. NO. 1.5e+03;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GSILOGS 6  
 |||  
 Db 1347 gsisigs 1352

## RESULT 15

AA057055  
 ID AA057055 standard; peptide: 21 AA.

XX AA057055;

XX 15-MAR-1995 (first entry)

XX HTLV peptide.

XX HTLV1; HTLV2; HTLV-I; HTLV-II; diagnostic.

XX HTLV.

XX MO9418322-A.

XX 18-AUG-1994.

XX 02-FEB-1994; 94MO-US01170.

XX 05-FEB-1993; 99US-0014153.

XX (GENE-) GENELABS TECHNOLOGIES INC.

XX (STRD ) UNIV IELAND STANFORD JUNIOR.

XX Fong SKH, Goh C, Hadlock KG;

XX WPI; 1994-279743/34.

XX HTLV-I and -II peptide(s) and kits contg. them - used to diagnose

XX infection and discriminate between HTLV-I and -II infection

XX Disclosure: Page 74; 100pp; English.

XX Novel HTLV1 and HTLV2 peptides are disclosed for use in diagnostic

XX assays for detecting HTLV1 and HTLV2 infection in human sera. 2

XX Peptides of the invention are given in AA057055-56.

XX Sequence 21 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX EP1033401-A2.

XX 06-SEP-2000.

XX 21-FEB-2000; 2000EP-0200610.

XX 26-FEB-1999; 99US-0122487.

XX (CEST ) GENSET.

XX Dumas Milne Edwards J, Duclert A, Giordano J;

XX WPI; 2000-500381/45.

XX M-PSDB; AAC01215.

XX New nucleic acid that is a 5' expressed sequence tag (5' EST) for

XX obtaining cDNAs and genomic DNAs that correspond to 5'ESTs and for

XX diagnostic, forensic, gene therapy and chromosome mapping procedures -

XX Claim 13; SEQ ID 5290; 71pp + CD-ROM; English.

XX The present sequence is a polypeptide encoded by one of a large number

XX of 5' ESTs derived from mRNAs encoding secreted proteins. The 5' ESTs

XX were prepared from total human RNAs or poly(A)<sup>+</sup> RNAs derived from 30

XX different tissues. EST sequences usually correspond mainly to the 3'

XX untranslated region (UTR) of the mRNA because they are often obtained

XX from oligo-dT primed cDNA libraries. Such ESTs are not well suited for

XX those cases where longer cDNA sequences have been obtained. The full 5'

XX UTR is rarely included. 5' ESTs are derived from mRNAs with intact 5'

XX ends and can therefore be used to obtain full length cDNAs and genomic

XX DNAs. 5' ESTs are also used in diagnostic, forensic, gene therapy and

XX regulatory sequences and to design expression and secretion vectors.

XX Sequence 97 AA;

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

Query Match 93.3%; Score 28; DB 21; Length 97;  
 Best Local Similarity 83.3%; Pred. NO. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSILOGS 6  
 |||  
 Db 34 gsisigs 39

## RESULT 17

AA074477  
 ID AA074477 standard; Protein; 119 AA.

XX AA074477;

XX 03-SEP-2001 (first entry)

XX Human colon cancer antigen protein SEQ ID NO:5241.

XX Human; colon cancer; colon cancer antigen; diagnosis; detection;

XX colorectal carcinoma.

XX Homo sapiens.

XX W0200122920-A2.

XX 05-APR-2001.

XX 28-SEP-2000; 2000MO-US26524.

XX 29-SEP-1999; 99US-0157137.

XX 03-NOV-1999; 99US-0163280.

XX (HMAN-) HUMAN GENOME SCI INC.

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX



XX Sequence 633 AA:  
SQ  
Query Match 93.3% Score 28: DB 19: Length 633:  
Best Local Similarity 83.3% Pred. No. 1.4e+03:  
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
OR 1 GSI/GGS 6  
DB 349 991995 354  
RESULT 20  
AAG50497 standard; protein: 1892 AA.  
XX AAG50497  
XX AAG50497:  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 64001.  
XX DE  
XX Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
XX termination sequence.  
XX Arabidopsis thaliana.  
XX EP103405-A2.  
XX 06-SEP-2000.  
XX 25-FEB-2000; 2000EP-0301439.  
XX 25-FEB-1999; 9905-0121835.  
PR 05-MAR-1999; 9905-0123180.  
PR 09-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 28-MAR-1999; 9905-0126284.  
PR 29-MAR-1999; 9905-0127462.  
PR 01-APR-1999; 9905-0128234.  
PR 06-APR-1999; 9905-0128714.  
PR 08-APR-1999; 9905-0129845.  
PR 16-APR-1999; 9905-0130077.  
PR 19-APR-1999; 9905-0130449.  
PR 21-APR-1999; 9905-0130510.  
PR 23-APR-1999; 9905-0131445.  
PR 28-APR-1999; 9905-0132408.  
PR 30-APR-1999; 9905-0132407.  
PR 30-APR-1999; 9905-0132484.  
PR 04-MAY-1999; 9905-0132485.  
PR 05-MAY-1999; 9905-0132486.  
PR 06-MAY-1999; 9905-0132487.  
PR 07-MAY-1999; 9905-0132486.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 14-MAY-1999; 9905-0134370.  
PR 19-MAY-1999; 9905-0134368.  
PR 20-MAY-1999; 9905-0135124.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 25-MAY-1999; 9905-0136392.  
PR 26-MAY-1999; 9905-0136782.  
PR 02-JUN-1999; 9905-0137252.  
PR 03-JUN-1999; 9905-0137252.  
PR 04-JUN-1999; 9905-0137502.

PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0138119.  
PR 16-JUN-1999; 9905-0139452.  
PR 18-JUN-1999; 9905-0139452.  
PR 17-JUN-1999; 9905-0139452.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139459.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 18-JUN-1999; 9905-0139750.  
PR 18-JUN-1999; 9905-0139753.  
PR 18-JUN-1999; 9905-0139817.  
PR 21-JUN-1999; 9905-0139817.  
PR 23-JUN-1999; 9905-0140523.  
PR 23-JUN-1999; 9905-0140525.  
PR 24-JUN-1999; 9905-0140595.  
PR 28-JUN-1999; 9905-0140823.  
PR 29-JUN-1999; 9905-0140991.  
PR 30-JUN-1999; 9905-0141287.  
PR 31-JUN-1999; 9905-0141582.  
PR 01-JUL-1999; 9905-0141584.  
PR 02-JUL-1999; 9905-0142035.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142803.  
PR 09-JUL-1999; 9905-0142920.  
PR 12-JUL-1999; 9905-0142977.  
PR 13-JUL-1999; 9905-0143542.  
PR 14-JUL-1999; 9905-0144005.  
PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144335.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144332.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144332.  
PR 20-JUL-1999; 9905-0144332.  
PR 20-JUL-1999; 9905-0144632.  
PR 20-JUL-1999; 9905-0144632.  
PR 21-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0144884.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145086.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-015152.  
PR 23-JUL-1999; 9905-015152.  
PR 23-JUL-1999; 9905-0152218.  
PR 23-JUL-1999; 9905-0152218.  
PR 26-JUL-1999; 9905-0152274.  
PR 26-JUL-1999; 9905-0152115.  
PR 27-JUL-1999; 9905-0145918.  
PR 27-JUL-1999; 9905-0145919.  
PR 28-JUL-1999; 9905-0145919.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146388.  
PR 03-AUG-1999; 9905-0147389.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 05-AUG-1999; 9905-0147260.  
PR 06-AUG-1999; 9905-0147303.

PR 06-AUG-1999: 99US-0147415  
 PR 09-AUG-1999: 99US-0147493  
 PR 09-AUG-1999: 99US-0147935  
 PR 10-AUG-1999: 99US-0148171  
 PR 11-AUG-1999: 99US-0148319  
 PR 12-AUG-1999: 99US-0148341  
 PR 13-AUG-1999: 99US-0148651  
 PR 14-AUG-1999: 99US-0148813  
 PR 16-AUG-1999: 99US-0149368  
 PR 17-AUG-1999: 99US-0149175  
 PR 18-AUG-1999: 99US-0149426  
 PR 20-AUG-1999: 99US-0149722  
 PR 20-AUG-1999: 99US-0149723  
 PR 20-AUG-1999: 99US-0149829  
 PR 23-AUG-1999: 99US-0149930  
 PR 25-AUG-1999: 99US-0150566  
 PR 26-AUG-1999: 99US-0150884  
 PR 27-AUG-1999: 99US-0151065  
 PR 27-AUG-1999: 99US-0151066  
 PR 27-AUG-1999: 99US-0151080  
 PR 31-AUG-1999: 99US-0151438  
 PR 01-SEP-1999: 99US-0151930  
 PR 07-SEP-1999: 99US-0152363  
 PR 10-SEP-1999: 99US-0153070  
 PR 11-SEP-1999: 99US-0153758  
 PR 15-SEP-1999: 99US-0154018  
 PR 15-SEP-1999: 99US-0154779  
 PR 23-SEP-1999: 99US-0155139  
 PR 23-SEP-1999: 99US-0155486  
 PR 24-SEP-1999: 99US-0155659  
 PR 26-SEP-1999: 99US-0156458  
 PR 29-SEP-1999: 99US-0156596  
 PR 04-OCT-1999: 99US-0157217  
 PR 06-OCT-1999: 99US-0157665  
 PR 07-OCT-1999: 99US-0158029  
 PR 08-OCT-1999: 99US-0158232  
 PR 12-OCT-1999: 99US-0158369  
 PR 13-OCT-1999: 99US-0159293  
 PR 13-OCT-1999: 99US-0159294  
 PR 13-OCT-1999: 99US-0159298  
 PR 14-OCT-1999: 99US-0159325  
 PR 14-OCT-1999: 99US-0159330  
 PR 14-OCT-1999: 99US-0159331  
 PR 14-OCT-1999: 99US-0159637  
 PR 14-OCT-1999: 99US-0159638  
 PR 18-OCT-1999: 99US-0159584  
 PR 18-OCT-1999: 99US-0159585  
 PR 21-OCT-1999: 99US-0160767  
 PR 21-OCT-1999: 99US-0160768  
 PR 21-OCT-1999: 99US-0160770  
 PR 21-OCT-1999: 99US-0160814  
 PR 21-OCT-1999: 99US-0160815  
 PR 22-OCT-1999: 99US-0160980  
 PR 22-OCT-1999: 99US-0160981  
 PR 23-OCT-1999: 99US-0161081  
 PR 25-OCT-1999: 99US-0161404  
 PR 25-OCT-1999: 99US-0161405  
 PR 25-OCT-1999: 99US-0161406  
 PR 26-OCT-1999: 99US-0161359  
 PR 26-OCT-1999: 99US-0161360  
 PR 26-OCT-1999: 99US-0161361  
 PR 28-OCT-1999: 99US-0161592  
 PR 28-OCT-1999: 99US-0161593  
 PR 29-OCT-1999: 99US-0162142

Query Match 93.3% Score 28: DB 21: Length 1892:  
 Best Local Similarity 83.3% Freq. NO. 44e+03:  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Oy 1 GSIGES 6  
 11:111  
 DB 650 gsl:igs 655  
 RESULT 21  
 AAC50496  
 ID AAC50496 standard: Protein: 1940 AA.  
 AC AAC50496:  
 DT 18-OCT-2000 (first entry)  
 XX  
 XX Arabidopsis thaliana protein fragment SPQ ID NO: 64000.  
 KW Protein identification: signal transduction pathway; metabolic  
 KW hydrolisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 KW  
 XX Arabidopsis thaliana.  
 XX  
 PN EPI033405-A2.  
 PD 06-SEP-2000.  
 XX  
 XX 25-FEB-2000: 2000EP-0301439.  
 XX  
 XX 25-FEB-1999: 99US-0121825  
 PR 05-MAR-1999: 99US-0123180  
 PR 09-MAR-1999: 99US-0123548  
 PR 23-MAR-1999: 99US-0125788  
 PR 25-MAR-1999: 99US-0126264  
 PR 29-MAR-1999: 99US-0126789  
 PR 01-APR-1999: 99US-0126934  
 PR 01-APR-1999: 99US-0126934  
 PR 08-APR-1999: 99US-0128714  
 PR 16-APR-1999: 99US-0129845  
 PR 19-APR-1999: 99US-0130077  
 PR 21-APR-1999: 99US-0130449  
 PR 23-APR-1999: 99US-0130510  
 PR 23-APR-1999: 99US-0130640  
 PR 23-APR-1999: 99US-0130640  
 PR 30-APR-1999: 99US-0132048  
 PR 04-MAY-1999: 99US-0132407  
 PR 04-MAY-1999: 99US-0132484  
 PR 05-MAY-1999: 99US-0132485  
 PR 06-MAY-1999: 99US-0132486  
 PR 06-MAY-1999: 99US-0132487  
 PR 11-MAY-1999: 99US-0134456  
 PR 14-MAY-1999: 99US-0134219  
 PR 14-MAY-1999: 99US-0134221  
 PR 14-MAY-1999: 99US-0134370  
 PR 18-MAY-1999: 99US-0134768  
 PR 18-MAY-1999: 99US-0134769  
 PR 23-MAY-1999: 99US-0135124  
 PR 24-MAY-1999: 99US-0135253  
 PR 24-MAY-1999: 99US-0135529  
 PR 25-MAY-1999: 99US-0136021  
 PR 27-MAY-1999: 99US-0136392  
 PR 28-MAY-1999: 99US-0136782  
 PR 01-JUN-1999: 99US-0137522  
 PR 04-JUN-1999: 99US-0137724  
 PR 07-JUN-1999: 99US-0137724  
 PR 08-JUN-1999: 99US-0138094  
 PR 10-JUN-1999: 99US-0138540  
 PR 10-JUN-1999: 99US-0138847  
 PR 14-JUN-1999: 99US-0139119  
 PR 16-JUN-1999: 99US-0139432  
 PR 16-JUN-1999: 99US-0139433

[illegible]

AA	AG50495	standard: Protein: 1944 AA.
ID	AG50495	
AC	AG50495:	
XX		
XX	18-OCT-2000 (first entry)	
XX		
XX	Arabidopsis thaliana protein fragment SEQ ID NO: 63999.	
DE		
XX		
KM	Protein Identification: signal transduction pathway; metabolic pathway;	
KM	hybridization assay; genetic mapping; gene expression control; promoter	
KM	termination sequence.	
XX		
XX	Arabidopsis thaliana.	
XX	EP103405-AZ.	
PD	06-SEP-2000.	
XX		
PE	25-FEB-2000; 2000EP-0301439.	
XX		
XX	25-SEP-1999.	9905-0121825.
PR	05-MAR-1999.	9905-0123160.
PR	09-MAR-1999.	9905-0123548.
PR	23-MAR-1999.	9905-0125788.
PR	25-MAR-1999.	9905-0126264.
PR	29-MAR-1999.	9905-0126785.
PR	01-APR-1999.	9905-0127462.
PR	08-APR-1999.	9905-0128234.
PR	08-APR-1999.	9905-0129645.
PR	15-APR-1999.	9905-0130677.
PR	21-APR-1999.	9905-0130449.
PR	23-APR-1999.	9905-0130510.
PR	23-APR-1999.	9905-0130891.
PR	28-APR-1999.	9905-0131449.
PR	30-APR-1999.	9905-0132065.
PR	04-MAY-1999.	9905-0132484.
PR	05-MAY-1999.	9905-0132485.
PR	06-MAY-1999.	9905-0132486.
PR	06-MAY-1999.	9905-0132487.
PR	07-MAY-1999.	9905-0132863.
PR	11-MAY-1999.	9905-0134256.
PR	14-MAY-1999.	9905-0134218.
PR	14-MAY-1999.	9905-0134421.
PR	14-MAY-1999.	9905-0134470.
PR	18-MAY-1999.	9905-0134768.
PR	19-MAY-1999.	9905-0134941.
PR	20-MAY-1999.	9905-0135124.
PR	21-MAY-1999.	9905-0135353.
PR	25-MAY-1999.	9905-0135823.
PR	25-MAY-1999.	9905-0136393.
PR	27-MAY-1999.	9905-0136782.
PR	03-JUN-1999.	9905-0137222.
PR	03-JUN-1999.	9905-0137528.
PR	04-JUN-1999.	9905-0137502.
PR	07-JUN-1999.	9905-0137724.
PR	08-JUN-1999.	9905-0138094.
PR	10-JUN-1999.	9905-0138843.
PR	10-JUN-1999.	9905-0139115.
PR	16-JUN-1999.	9905-0139452.
PR	16-JUN-1999.	9905-0139453.
PR	17-JUN-1999.	9905-0139457.
PR	18-JUN-1999.	9905-0139454.
PR	18-JUN-1999.	9905-0139455.
PR	18-JUN-1999.	9905-0139456.
PR	18-JUN-1999.	9905-0139457.
PR	18-JUN-1999.	9905-0139458.
PR	18-JUN-1999.	9905-0139459.
PR	18-JUN-1999.	9905-0139460.
PR	18-JUN-1999.	9905-0139461.
PR	18-JUN-1999.	9905-0139462.
PR	18-JUN-1999.	9905-0139463.
PR	18-JUN-1999.	9905-0139464.
PR	18-JUN-1999.	9905-0139465.
PR	18-JUN-1999.	9905-0139466.
PR	18-JUN-1999.	9905-0139467.
PR	18-JUN-1999.	9905-0139468.
PR	18-JUN-1999.	9905-0139469.
PR	18-JUN-1999.	9905-0139470.
PR	18-JUN-1999.	9905-0139471.
PR	18-JUN-1999.	9905-0139472.
PR	18-JUN-1999.	9905-0139473.
PR	18-JUN-1999.	9905-0139474.
PR	18-JUN-1999.	9905-0139475.
PR	18-JUN-1999.	9905-0139476.
PR	18-JUN-1999.	9905-0139477.
PR	18-JUN-1999.	9905-0139478.
PR	18-JUN-1999.	9905-0139479.
PR	18-JUN-1999.	9905-0139480.
PR	18-JUN-1999.	9905-0139481.
PR	18-JUN-1999.	9905-0139482.
PR	18-JUN-1999.	9905-0139483.
PR	18-JUN-1999.	9905-0139484.
PR	18-JUN-1999.	9905-0139485.
PR	18-JUN-1999.	9905-0139486.
PR	18-JUN-1999.	9905-0139487.
PR	18-JUN-1999.	9905-0139488.
PR	18-JUN-1999.	9905-0139489.
PR	18-JUN-1999.	9905-0139490.
PR	18-JUN-1999.	9905-0139491.
PR	18-JUN-1999.	9905-0139492.
PR	18-JUN-1999.	9905-0139493.
PR	18-JUN-1999.	9905-0139494.
PR	18-JUN-1999.	9905-0

[illegible]

```

PR 23-AUG-1999; 9905-0145902.
PR 23-AUG-1999; 9905-0145930.
PR 23-AUG-1999; 9905-0150586.
PR 27-AUG-1999; 9905-0151085.
PR 27-AUG-1999; 9905-0151086.
PR 27-AUG-1999; 9905-0151086.
PR 30-AUG-1999; 9905-0151303.
PR 31-AUG-1999; 9905-0151438.
PR 01-SEP-1999; 9905-0151930.
PR 10-SEP-1999; 9905-0153053.
PR 13-SEP-1999; 9905-0153758.
PR 15-SEP-1999; 9905-0154038.
PR 16-SEP-1999; 9905-0154039.
PR 20-SEP-1999; 9905-0154779.
PR 22-SEP-1999; 9905-0155139.
PR 23-SEP-1999; 9905-0155486.
PR 24-SEP-1999; 9905-0155486.
PR 26-SEP-1999; 9905-0155486.
PR 29-SEP-1999; 9905-0155596.
PR 04-OCT-1999; 9905-0157117.
PR 05-OCT-1999; 9905-0157753.
PR 06-OCT-1999; 9905-0157865.
PR 07-OCT-1999; 9905-0158029.
PR 08-OCT-1999; 9905-0158422.
PR 11-OCT-1999; 9905-0158523.
PR 13-OCT-1999; 9905-0159294.
PR 13-OCT-1999; 9905-0159295.
PR 14-OCT-1999; 9905-0159329.
PR 14-OCT-1999; 9905-0159330.
PR 14-OCT-1999; 9905-0159331.
PR 14-OCT-1999; 9905-0159332.
PR 14-OCT-1999; 9905-0159584.
PR 18-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160741.
PR 21-OCT-1999; 9905-0160767.
PR 21-OCT-1999; 9905-0160768.
PR 21-OCT-1999; 9905-0160770.
PR 21-OCT-1999; 9905-0160814.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160980.
PR 22-OCT-1999; 9905-0160981.
PR 22-OCT-1999; 9905-0160989.
PR 25-OCT-1999; 9905-0161404.
PR 25-OCT-1999; 9905-0161405.
PR 25-OCT-1999; 9905-0161406.
PR 26-OCT-1999; 9905-0161329.
PR 26-OCT-1999; 9905-0161361.
PR 26-OCT-1999; 9905-0161361.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161992.
PR 28-OCT-1999; 9905-0161993.
PR 29-OCT-1999; 9905-0162142.

```

```

Query Match          93.38; Score 28; DB 21; Length 1944;
Best Local Similarity 83.38; Pred. No. 26+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1  GSLGGS 6
      ||:||||
Db      702  gslggs 707

```

```

RESULT 23
AAV04755
ID AAV04755 standard; Protein: 67 AA.
AC AAV04755;
XX
XX 06-JUL-1999 (first entry)
XX

```

```

DE Mycobacterium species protein sequence 244.
XX
XX Secreted protein; Mycobacterium; primer; PCR; amplification; probe;
KM hybridisation; detection; vaccine; immunisation; infection.
XX
XX Mycobacterium sp.
OS
XX WC9090186-A2.
XX
XX 25-FEB-1999.
PD
XX 14-AUG-1998; 98MO-FR01813.
XX
XX 11-SEP-1997; 97FR-0011325.
PR 14-AUG-1997; 97FR-0010404.
XX
XX (IMSP ) INST PASTEUR.
FA
XX Glucanase B; Lim EM, Pellic V, Portnoi D, Goguet de la Salmoniere Y;
PI
XX Guineau A;
XX
XX MPI: 1999-181045/15.
DR
XX N-PSDB: AAX34007.
XX
XX
XX Mycobacterial DNA vectors containing reporter constructs - for
PT Identifying coding or promoter sequences involved in
PT Infection-associated protein expression
XX
XX Claim 32, Fig 2, 309pp; French.
XX
XX Sequences AAV04742-Y05000 and AAV07201-Y07204 represent secreted
CC proteins from various Mycobacterium species microorganisms. The
CC encoding nucleotide sequences can be used as primers and probes for
CC methods for detecting and identifying mycobacteria, especially belonging
CC to the M. tuberculosis complex. The encoded proteins can be used in
XX vaccines for immunisation against a bacterial or viral infection.
XX
XX Sequence 67 AA:
SQ

```

```

Query Match          90.04; Score 27; DB 20; Length 67;
Best Local Similarity 83.34; Pred. No. 26+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy      1  GSLGGS 6
      ||:||||
Db      59  gslggs 64

```

```

RESULT 24
AAZ27140
ID AAZ27140 standard; Protein: 91 AA.
XX
XX AAZ27140;
XX

```

```

DT 17-OCT-2000 (first entry)

```

```

DE Zea mays protein fragment SEQ ID NO: 31861.

```

```

XX Protein identification; signal transduction pathway; metabolic pathway;
KM hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence; corn.
XX
XX Zea mays subsp. mays.
OS
XX EP1033405-A2.
XX
XX 06-SEP-2000.
PD
XX 25-FEB-2000; 2000EP-0301439.
XX
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.

```

PR 03-MAR-1999; 9905-0123548.  
PR 23-MAR-1999; 9905-0125788.  
PR 28-MAR-1999; 9905-0126284.  
PR 01-APR-1999; 9905-0127184.  
PR 01-APR-1999; 9905-0127462.  
PR 06-APR-1999; 9905-0128234.  
PR 08-APR-1999; 9905-0128714.  
PR 16-APR-1999; 9905-0129845.  
PR 19-APR-1999; 9905-0130077.  
PR 21-APR-1999; 9905-0130449.  
PR 23-APR-1999; 9905-0131141.  
PR 28-APR-1999; 9905-0131449.  
PR 30-APR-1999; 9905-0132048.  
PR 30-APR-1999; 9905-0132407.  
PR 04-MAY-1999; 9905-0132484.  
PR 05-MAY-1999; 9905-0132485.  
PR 06-MAY-1999; 9905-0132486.  
PR 07-MAY-1999; 9905-0132657.  
PR 11-MAY-1999; 9905-0134256.  
PR 14-MAY-1999; 9905-0134218.  
PR 14-MAY-1999; 9905-0134219.  
PR 14-MAY-1999; 9905-0134221.  
PR 14-MAY-1999; 9905-0134370.  
PR 19-MAY-1999; 9905-0134768.  
PR 20-MAY-1999; 9905-0135121.  
PR 21-MAY-1999; 9905-0135353.  
PR 24-MAY-1999; 9905-0135629.  
PR 25-MAY-1999; 9905-0136021.  
PR 27-MAY-1999; 9905-0136192.  
PR 28-MAY-1999; 9905-0136782.  
PR 03-JUN-1999; 9905-0137522.  
PR 03-JUN-1999; 9905-0137502.  
PR 07-JUN-1999; 9905-0137724.  
PR 08-JUN-1999; 9905-0138094.  
PR 10-JUN-1999; 9905-0138540.  
PR 10-JUN-1999; 9905-0138847.  
PR 14-JUN-1999; 9905-0139119.  
PR 16-JUN-1999; 9905-0139453.  
PR 17-JUN-1999; 9905-0139492.  
PR 18-JUN-1999; 9905-0139454.  
PR 18-JUN-1999; 9905-0139455.  
PR 18-JUN-1999; 9905-0139456.  
PR 18-JUN-1999; 9905-0139457.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139458.  
PR 18-JUN-1999; 9905-0139460.  
PR 18-JUN-1999; 9905-0139461.  
PR 18-JUN-1999; 9905-0139462.  
PR 18-JUN-1999; 9905-0139463.  
PR 23-JUN-1999; 9905-0139899.  
PR 23-JUN-1999; 9905-0140353.  
PR 23-JUN-1999; 9905-0140354.  
PR 24-JUN-1999; 9905-0140695.  
PR 28-JUN-1999; 9905-0140823.  
PR 28-JUN-1999; 9905-0140991.  
PR 01-JUL-1999; 9905-0141187.  
PR 01-JUL-1999; 9905-0142154.  
PR 02-JUL-1999; 9905-0142055.  
PR 06-JUL-1999; 9905-0142390.  
PR 08-JUL-1999; 9905-0142603.  
PR 09-JUL-1999; 9905-0142920.  
PR 11-JUL-1999; 9905-0142877.  
PR 14-JUL-1999; 9905-0143624.  
PR 14-JUL-1999; 9905-0143624.

PR 15-JUL-1999; 9905-0144005.  
PR 16-JUL-1999; 9905-0144085.  
PR 16-JUL-1999; 9905-0144086.  
PR 19-JUL-1999; 9905-0144325.  
PR 19-JUL-1999; 9905-0144331.  
PR 19-JUL-1999; 9905-0144333.  
PR 19-JUL-1999; 9905-0144334.  
PR 19-JUL-1999; 9905-0144335.  
PR 20-JUL-1999; 9905-0144352.  
PR 20-JUL-1999; 9905-0144352.  
PR 21-JUL-1999; 9905-0144894.  
PR 21-JUL-1999; 9905-0145086.  
PR 21-JUL-1999; 9905-0145088.  
PR 22-JUL-1999; 9905-0145085.  
PR 22-JUL-1999; 9905-0145087.  
PR 22-JUL-1999; 9905-0145089.  
PR 22-JUL-1999; 9905-0145192.  
PR 23-JUL-1999; 9905-0145145.  
PR 23-JUL-1999; 9905-0145245.  
PR 23-JUL-1999; 9905-0145276.  
PR 26-JUL-1999; 9905-0145276.  
PR 27-JUL-1999; 9905-0145313.  
PR 27-JUL-1999; 9905-0145318.  
PR 27-JUL-1999; 9905-0145319.  
PR 28-JUL-1999; 9905-0145351.  
PR 02-AUG-1999; 9905-0146386.  
PR 02-AUG-1999; 9905-0146389.  
PR 04-AUG-1999; 9905-0147038.  
PR 04-AUG-1999; 9905-0147204.  
PR 04-AUG-1999; 9905-0147302.  
PR 05-AUG-1999; 9905-0147192.  
PR 06-AUG-1999; 9905-0147300.  
PR 06-AUG-1999; 9905-0147316.  
PR 09-AUG-1999; 9905-0147493.  
PR 09-AUG-1999; 9905-0147535.  
PR 10-AUG-1999; 9905-0148171.  
PR 11-AUG-1999; 9905-0148319.  
PR 11-AUG-1999; 9905-0148341.  
PR 11-AUG-1999; 9905-0148682.  
PR 11-AUG-1999; 9905-0148683.  
PR 17-AUG-1999; 9905-0149368.  
PR 17-AUG-1999; 9905-0149375.  
PR 18-AUG-1999; 9905-0149426.  
PR 20-AUG-1999; 9905-0149722.  
PR 20-AUG-1999; 9905-0149723.  
PR 21-AUG-1999; 9905-0149829.  
PR 21-AUG-1999; 9905-0149830.  
PR 21-AUG-1999; 9905-0149930.  
PR 25-AUG-1999; 9905-0150566.  
PR 26-AUG-1999; 9905-0150884.  
PR 27-AUG-1999; 9905-0151665.  
PR 27-AUG-1999; 9905-0151666.  
PR 27-AUG-1999; 9905-0151680.  
PR 31-AUG-1999; 9905-0151438.  
PR 01-SEP-1999; 9905-0151930.  
PR 07-SEP-1999; 9905-0152363.  
PR 10-SEP-1999; 9905-0153070.  
PR 11-SEP-1999; 9905-0153758.  
PR 11-SEP-1999; 9905-0154018.  
PR 16-SEP-1999; 9905-0154939.  
PR 20-SEP-1999; 9905-0154939.  
PR 23-SEP-1999; 9905-0155139.  
PR 23-SEP-1999; 9905-0155486.  
PR 24-SEP-1999; 9905-0155559.  
PR 26-SEP-1999; 9905-0156596.  
PR 04-OCT-1999; 9905-0157117.  
PR 04-OCT-1999; 9905-0157753.





XX Rosenthal A, Specht T, Hinzmann B, Schmitt A, Pilarsky C, Dahl E;  
 XX MPI: 1999-591957/51.  
 XX N-PSDB: AA41953.  
 PT New nucleic acid sequences expressed in uterine cancer tissues, and  
 PT derived polypeptides, for treatment of uterine and endometrial cancer  
 PT and identification of therapeutic agents -  
 PS Claim 23: Page 290; 444pp: German.  
 XX This invention describes novel human nucleic acid (cDNA) sequences (A),  
 XX that are highly expressed in uterine tumour tissue and which have  
 XX anticancer and cytostatic activity. (A) are used (1) for recombinant  
 XX expression of polypeptides (B) and (11) to isolate complete genes. (B)  
 XX are used (1) to identify agents suitable for treatment of uterine or  
 XX endometrial cancer; (11) directly for treating these forms of cancer  
 XX (including expression from gene therapy vectors) and (111) for  
 XX generation of specific antibodies. (A) are identified by assembling ESTs  
 XX of expressed sequence tags (ESTs) and sequence information of expression  
 XX of expression patterns. This allows a significantly longer fragment of  
 XX the gene to be revealed, so should reduce the number of failures  
 XX associated with the fact that ESTs from different libraries may represent  
 XX different parts of the same unknown gene, distorting the estimated  
 XX frequency of occurrence in a particular tissue. AA59941-Y60328 represent  
 XX protein fragments encoded by the human endometrium tumour cDNA library  
 XX derived EST fragments represented in AA41981-242121.  
 SO Sequence 126 AA:

Query Match 90.0% Score 27: DB 20: Length 126:  
 Best Local Similarity 83.3% Pred No. 4e+02: 0: Gaps 0:  
 Matches 5: Conservative 1: Mismatches 0: Indels 0:  
 QY 1 GSLGGS 6  
 DB 52 gtlggs 57

RESULT 27  
 AA59504  
 IAA59504 standard: Protein: 126 AA.  
 AC AA59504:  
 XX 26-JUN-2001 (first entry)  
 D7 XX  
 XX Human protein sequence SEQ ID NO:16665.  
 DE XX  
 XX Human: primer: detection: diagnosis: antisense therapy: gene therapy.  
 KW XX  
 XX Homo sapiens.  
 OS XX  
 XX EPI074617-A2.  
 PN XX  
 XX 07-FEB-2001.  
 XX XX  
 XX 28-JUL-2000: 2000EP-0116126.  
 PE XX  
 XX 29-JUL-1999: 99JP-0248036.  
 PR 27-AUG-1999: 99JP-0300253.  
 PR 11-JUN-2000: 2000JP-0118776.  
 PR 02-MAR-2000: 2000JP-0283176.  
 PR 03-JUN-2000: 2000JP-0241893.  
 XX XX  
 XX (HELI-) HELIX RES INST.  
 PA  
 PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;  
 PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;  
 XX WPI: 2001-318749/34.

XX Primer sets for synthesizing polynucleotides, particularly the 5602  
 PT full-length cDNAs defined in the specification, and for the detection  
 PT and/or diagnosis of the abnormality of the proteins encoded by the  
 PT full-length cDNAs -  
 XX  
 XX Claim 8: SEQ ID 16665: 2837bp + CD ROW: English.  
 PS  
 XX The present invention describes primer sets for synthesizing 5602  
 XX full-length cDNAs defined in the specification. Where a primer set  
 XX comprises: (a) an oligo- or primer and an oligonucleotide complementary  
 XX to the 5602 full-length cDNA sequence defined in the specification, where the  
 XX oligonucleotide comprises at least 15 nucleotides; or (b) a combination  
 XX of an oligonucleotide comprising a sequence complementary to the  
 XX complementary strand of a polynucleotide which comprises a 5'-end  
 XX sequence and an oligonucleotide comprising a sequence complementary to a  
 XX polynucleotide which comprises at least 15 nucleotides, where the  
 XX oligonucleotide comprises a 3'-end sequence, where the combination of  
 XX the 3'-end sequence and sequence can be used in antisense therapy and  
 XX in gene therapy. The primers are useful for synthesizing polynucleotides,  
 XX particularly full-length cDNAs. The primers are also useful for the  
 XX detection and/or diagnosis of the abnormality of the proteins encoded by  
 XX the full-length cDNAs. The primers allow obtaining of the full-length  
 XX cDNAs easily without any specialised methods. AA403166 to AA415628 and  
 XX AA415633 to AA418742 represent human cDNA sequences; AA19323 to AA19332  
 XX represent oligonucleotides, all of which are used in the exemplification  
 XX of the present invention.  
 SO Sequence 126 AA:

Query Match 90.0% Score 27: DB 22: Length 126:  
 Best Local Similarity 83.3% Pred No. 4e+02: 0: Gaps 0:  
 Matches 5: Conservative 1: Mismatches 0: Indels 0:  
 QY 1 GSLGGS 6  
 DB 101 gswggs 106

RESULT 28  
 AA453009  
 ID AA453009 standard: Protein: 163 AA.  
 XX  
 AC AA453009:  
 XX 03-AUG-1998 (first entry)  
 D7 XX  
 XX Mus musculus I-mfc protein.  
 DE XX  
 XX I-mfc: Inhibitor of MYD family; treatment: diagnosis: myogenesis;  
 KW defects: abnormal development: disease: cleft/dorsal dysplasia;  
 KW CDD: Fibroadenoma; muscle tissue cancer.  
 KM  
 XX Mus musculus.  
 OS  
 XX M09808860-AL.  
 PN  
 XX 05-MAR-1998.  
 PD  
 XX 21-AUG-1997: 97MO-USJ4780.  
 PE  
 XX 27-AUG-1996: 96US-0704931.  
 XX XX  
 XX (HUTC-) HUTCHINSON CANCER RES CENT: FRED.  
 PA  
 PI Chen CMA, Groudine M, Kraut N, Weintraub H;  
 PI WPI: 1996-179377/16.  
 XX N-PSDB: AA42184.

PR 28-MAY-1999; 9905-0136762  
PR 01-JUN-1999; 9906-0137222

PR 05-JUN-1999: 99US-0137528.  
PR 04-JUN-1999: 99US-0137502.  
PR 03-JUN-1999: 99US-0137724.  
PR 02-JUN-1999: 99US-0137724.  
PR 10-JUN-1999: 99US-0138540.  
PR 10-JUN-1999: 99US-0138847.  
PR 14-JUN-1999: 99US-0139119.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139453.  
PR 16-JUN-1999: 99US-0139452.  
PR 16-JUN-1999: 99US-0139452.  
PR 18-JUN-1999: 99US-0139455.  
PR 18-JUN-1999: 99US-0139456.  
PR 18-JUN-1999: 99US-0139457.  
PR 18-JUN-1999: 99US-0139458.  
PR 18-JUN-1999: 99US-0139459.  
PR 18-JUN-1999: 99US-0139460.  
PR 18-JUN-1999: 99US-0139461.  
PR 18-JUN-1999: 99US-0139463.  
PR 18-JUN-1999: 99US-0139750.  
PR 18-JUN-1999: 99US-0139763.  
PR 18-JUN-1999: 99US-0139817.  
PR 21-JUN-1999: 99US-0139817.  
PR 22-JUN-1999: 99US-0139899.  
PR 22-JUN-1999: 99US-0140353.  
PR 23-JUN-1999: 99US-0140353.  
PR 24-JUN-1999: 99US-0140695.  
PR 28-JUN-1999: 99US-0140823.  
PR 29-JUN-1999: 99US-0140991.  
PR 30-JUN-1999: 99US-0141287.  
PR 01-JUL-1999: 99US-0141842.  
PR 01-JUL-1999: 99US-0142154.  
PR 02-JUL-1999: 99US-0142350.  
PR 06-JUL-1999: 99US-0142803.  
PR 08-JUL-1999: 99US-0142930.  
PR 12-JUL-1999: 99US-0142977.  
PR 13-JUL-1999: 99US-0143542.  
PR 15-JUL-1999: 99US-0143644.  
PR 15-JUL-1999: 99US-0144052.  
PR 16-JUL-1999: 99US-0144086.  
PR 16-JUL-1999: 99US-0144335.  
PR 19-JUL-1999: 99US-0144335.  
PR 19-JUL-1999: 99US-0144332.  
PR 19-JUL-1999: 99US-0144333.  
PR 19-JUL-1999: 99US-0144334.  
PR 19-JUL-1999: 99US-0144334.  
PR 20-JUL-1999: 99US-0144352.  
PR 20-JUL-1999: 99US-0144652.  
PR 20-JUL-1999: 99US-0144884.  
PR 21-JUL-1999: 99US-0144814.  
PR 21-JUL-1999: 99US-0145086.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145087.  
PR 22-JUL-1999: 99US-0145089.  
PR 22-JUL-1999: 99US-0145192.  
PR 23-JUL-1999: 99US-0145145.  
PR 23-JUL-1999: 99US-0145218.  
PR 23-JUL-1999: 99US-0145218.  
PR 26-JUL-1999: 99US-0145214.  
PR 27-JUL-1999: 99US-0145213.  
PR 27-JUL-1999: 99US-0145216.  
PR 27-JUL-1999: 99US-0145219.  
PR 28-JUL-1999: 99US-0145521.  
PR 28-JUL-1999: 99US-0145386.  
PR 02-AUG-1999: 99US-0146388.  
PR 02-AUG-1999: 99US-0146388.  
PR 03-AUG-1999: 99US-0146389.  
PR 04-AUG-1999: 99US-0147204.  
PR 04-AUG-1999: 99US-0147202.  
PR 05-AUG-1999: 99US-0147192.

PR 05-AUG-1999: 99US-0147260.  
PR 06-AUG-1999: 99US-0147303.  
PR 06-AUG-1999: 99US-0147416.  
PR 09-AUG-1999: 99US-0147493.  
PR 09-AUG-1999: 99US-0147935.  
PR 10-AUG-1999: 99US-0148171.  
PR 10-AUG-1999: 99US-0148319.  
PR 12-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148565.  
PR 13-AUG-1999: 99US-0148664.  
PR 16-AUG-1999: 99US-0149368.  
PR 17-AUG-1999: 99US-0149175.  
PR 18-AUG-1999: 99US-0149426.  
PR 20-AUG-1999: 99US-0149722.  
PR 20-AUG-1999: 99US-0149929.  
PR 23-AUG-1999: 99US-0149902.  
PR 23-AUG-1999: 99US-0149902.  
PR 25-AUG-1999: 99US-0150566.  
PR 25-AUG-1999: 99US-0150884.  
PR 26-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151065.  
PR 27-AUG-1999: 99US-0151080.  
PR 30-AUG-1999: 99US-0151303.  
PR 01-AUG-1999: 99US-0151338.  
PR 01-SEP-1999: 99US-0151330.  
PR 01-SEP-1999: 99US-0152633.  
PR 01-SEP-1999: 99US-0152720.  
PR 13-SEP-1999: 99US-0153070.  
PR 15-SEP-1999: 99US-0154018.  
PR 16-SEP-1999: 99US-0154039.  
PR 20-SEP-1999: 99US-0154779.  
PR 22-SEP-1999: 99US-0155139.  
PR 23-SEP-1999: 99US-0155486.  
PR 24-SEP-1999: 99US-0155859.  
PR 25-SEP-1999: 99US-0156358.  
PR 28-SEP-1999: 99US-0156358.  
PR 04-OCT-1999: 99US-0157717.  
PR 05-OCT-1999: 99US-0157753.  
PR 06-OCT-1999: 99US-0157665.  
PR 07-OCT-1999: 99US-0158029.  
PR 08-OCT-1999: 99US-0158232.  
PR 11-OCT-1999: 99US-0158359.  
PR 11-OCT-1999: 99US-0158594.  
PR 13-OCT-1999: 99US-0158594.  
PR 13-OCT-1999: 99US-0159293.  
PR 14-OCT-1999: 99US-0159329.  
PR 14-OCT-1999: 99US-0159330.  
PR 14-OCT-1999: 99US-0159331.  
PR 14-OCT-1999: 99US-0159652.  
PR 14-OCT-1999: 99US-0159652.  
PR 18-OCT-1999: 99US-0159584.  
PR 21-OCT-1999: 99US-0160741.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160767.  
PR 21-OCT-1999: 99US-0160814.  
PR 21-OCT-1999: 99US-0160814.  
PR 22-OCT-1999: 99US-0160980.  
PR 22-OCT-1999: 99US-0160981.  
PR 22-OCT-1999: 99US-0160989.  
PR 25-OCT-1999: 99US-0161404.  
PR 25-OCT-1999: 99US-0161405.  
PR 25-OCT-1999: 99US-0161406.  
PR 26-OCT-1999: 99US-0161329.  
PR 26-OCT-1999: 99US-0161361.  
PR 28-OCT-1999: 99US-0161920.  
PR 28-OCT-1999: 99US-0161992.  
PR 28-OCT-1999: 99US-0161993.  
PR 29-OCT-1999: 99US-0162142.

Query Match

90.0%, Score 27, DB 21, Length 188.



PR 04-JUN-1999: 99US-0137502.  
 PR 07-JUN-1999: 99US-0137724.  
 PR 08-JUN-1999: 99US-0138094.  
 PR 10-JUN-1999: 99US-0138540.  
 PR 10-JUN-1999: 99US-0138847.  
 PR 14-JUN-1999: 99US-0139415.  
 PR 15-JUN-1999: 99US-0139453.  
 PR 17-JUN-1999: 99US-0139492.  
 PR 18-JUN-1999: 99US-0139454.  
 PR 18-JUN-1999: 99US-0139455.  
 PR 18-JUN-1999: 99US-0139456.  
 PR 18-JUN-1999: 99US-0139457.  
 PR 18-JUN-1999: 99US-0139458.  
 PR 18-JUN-1999: 99US-0139459.  
 PR 18-JUN-1999: 99US-0139460.  
 PR 18-JUN-1999: 99US-0139461.  
 PR 18-JUN-1999: 99US-0139462.  
 PR 18-JUN-1999: 99US-0139463.  
 PR 18-JUN-1999: 99US-0139470.  
 PR 21-JUN-1999: 99US-0139817.  
 PR 22-JUN-1999: 99US-0139889.  
 PR 23-JUN-1999: 99US-0140353.  
 PR 23-JUN-1999: 99US-0140354.  
 PR 24-JUN-1999: 99US-0140695.  
 PR 28-JUN-1999: 99US-0140883.  
 PR 28-JUN-1999: 99US-0142870.  
 PR 28-JUN-1999: 99US-0143281.  
 PR 01-JUL-1999: 99US-0148482.  
 PR 02-JUL-1999: 99US-0142154.  
 PR 06-JUL-1999: 99US-0142055.  
 PR 06-JUL-1999: 99US-0143390.  
 PR 09-JUL-1999: 99US-0143803.  
 PR 09-JUL-1999: 99US-0143870.  
 PR 13-JUL-1999: 99US-0143542.  
 PR 14-JUL-1999: 99US-0143654.  
 PR 15-JUL-1999: 99US-0144005.  
 PR 16-JUL-1999: 99US-0144085.  
 PR 16-JUL-1999: 99US-0144086.  
 PR 19-JUL-1999: 99US-0144335.  
 PR 19-JUL-1999: 99US-0144336.  
 PR 19-JUL-1999: 99US-0144337.  
 PR 19-JUL-1999: 99US-0144338.  
 PR 19-JUL-1999: 99US-0144339.  
 PR 19-JUL-1999: 99US-0144340.  
 PR 19-JUL-1999: 99US-0144341.  
 PR 19-JUL-1999: 99US-0144342.  
 PR 19-JUL-1999: 99US-0144343.  
 PR 19-JUL-1999: 99US-0144344.  
 PR 19-JUL-1999: 99US-0144345.  
 PR 20-JUL-1999: 99US-0144352.  
 PR 20-JUL-1999: 99US-0144632.  
 PR 20-JUL-1999: 99US-0144894.  
 PR 21-JUL-1999: 99US-0145086.  
 PR 21-JUL-1999: 99US-0145088.  
 PR 22-JUL-1999: 99US-0145085.  
 PR 22-JUL-1999: 99US-0145087.  
 PR 22-JUL-1999: 99US-0145089.  
 PR 22-JUL-1999: 99US-0145132.  
 PR 22-JUL-1999: 99US-0145133.  
 PR 23-JUL-1999: 99US-0145224.  
 PR 23-JUL-1999: 99US-0145224.  
 PR 26-JUL-1999: 99US-0145276.  
 PR 27-JUL-1999: 99US-0145913.  
 PR 27-JUL-1999: 99US-0145918.  
 PR 28-JUL-1999: 99US-0145919.  
 PR 28-JUL-1999: 99US-0146381.  
 PR 02-AUG-1999: 99US-0146388.  
 PR 02-AUG-1999: 99US-0146389.  
 PR 03-AUG-1999: 99US-0147038.  
 PR 04-AUG-1999: 99US-0147204.  
 PR 04-AUG-1999: 99US-0147302.  
 PR 05-AUG-1999: 99US-0147152.  
 PR 05-AUG-1999: 99US-0147260.

PR 06-AUG-1999: 99US-0147703.  
 PR 06-AUG-1999: 99US-0147716.  
 PR 09-AUG-1999: 99US-0147923.  
 PR 09-AUG-1999: 99US-0147925.  
 PR 11-AUG-1999: 99US-0148310.  
 PR 11-AUG-1999: 99US-0148311.  
 PR 13-AUG-1999: 99US-0148345.  
 PR 13-AUG-1999: 99US-0148346.  
 PR 16-AUG-1999: 99US-0148684.  
 PR 17-AUG-1999: 99US-0149172.  
 PR 18-AUG-1999: 99US-0149172.  
 PR 20-AUG-1999: 99US-0149723.  
 PR 20-AUG-1999: 99US-0149729.  
 PR 23-AUG-1999: 99US-0149902.  
 PR 23-AUG-1999: 99US-0149902.  
 PR 23-AUG-1999: 99US-0149903.  
 PR 25-AUG-1999: 99US-0150366.  
 PR 25-AUG-1999: 99US-0150366.  
 PR 25-AUG-1999: 99US-0150884.  
 PR 27-AUG-1999: 99US-0153758.  
 PR 27-AUG-1999: 99US-0151066.  
 PR 30-AUG-1999: 99US-0151080.  
 PR 31-AUG-1999: 99US-0151083.  
 PR 01-SEP-1999: 99US-0151338.  
 PR 07-SEP-1999: 99US-0152263.  
 PR 10-SEP-1999: 99US-0152263.  
 PR 11-SEP-1999: 99US-0153758.  
 PR 15-SEP-1999: 99US-0154018.  
 PR 16-SEP-1999: 99US-0154039.  
 PR 20-SEP-1999: 99US-0154779.  
 PR 22-SEP-1999: 99US-0155139.  
 PR 23-SEP-1999: 99US-0155486.  
 PR 24-SEP-1999: 99US-0155486.  
 PR 28-SEP-1999: 99US-0156458.  
 PR 28-SEP-1999: 99US-0156596.  
 PR 04-OCT-1999: 99US-0157117.  
 PR 05-OCT-1999: 99US-0157753.  
 PR 06-OCT-1999: 99US-0157865.  
 PR 07-OCT-1999: 99US-0158029.  
 PR 08-OCT-1999: 99US-0158472.  
 PR 11-OCT-1999: 99US-0159233.  
 PR 13-OCT-1999: 99US-0159234.  
 PR 13-OCT-1999: 99US-0159235.  
 PR 14-OCT-1999: 99US-0159329.  
 PR 14-OCT-1999: 99US-0159330.  
 PR 14-OCT-1999: 99US-0159331.  
 PR 14-OCT-1999: 99US-0159332.  
 PR 14-OCT-1999: 99US-0159333.  
 PR 18-OCT-1999: 99US-0159584.  
 PR 21-OCT-1999: 99US-0160771.  
 PR 21-OCT-1999: 99US-0160767.  
 PR 21-OCT-1999: 99US-0160768.  
 PR 21-OCT-1999: 99US-0160770.  
 PR 21-OCT-1999: 99US-0160815.  
 PR 22-OCT-1999: 99US-0160980.  
 PR 22-OCT-1999: 99US-0160981.  
 PR 22-OCT-1999: 99US-0160989.  
 PR 25-OCT-1999: 99US-0161404.  
 PR 25-OCT-1999: 99US-0161405.  
 PR 25-OCT-1999: 99US-0161406.  
 PR 26-OCT-1999: 99US-0161380.  
 PR 26-OCT-1999: 99US-0161381.  
 PR 28-OCT-1999: 99US-0161993.  
 PR 28-OCT-1999: 99US-0161992.  
 PR 29-OCT-1999: 99US-0162112.

Query Match 90.0% Score 27: PR 21: Length 192:  
 Best Local Similarity 83.3% Pred. No. 6.2e+02:

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
 Oy 1 GSIGGS 6  
 Db 74 9a199s 79

## RESULT 33

AAW81735  
 ID AAW81735 standard; Protein: 196 AA.

AC AAW81735:  
 AA64370;

DE 27-JAN-1999 (first entry)

DE M. tuberculosis immunogenic polypeptide XP27 protein #2.

KW Tuberculosis; immunogenic; soluble; antigen; protective immunity; TB;  
 KW vaccine; pharmaceutical; infection; diagnosis.

DS Mycobacterium tuberculosis.

PN M09816646-A2.

PD 23-APR-1998.

PE 07-OCT-1997; 97MO-US18293.

PR 13-MAR-1997; 97US-0818112.

PR 11-OCT-1996; 96US-0730510.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;  
 PI Reed SC, Skeiky YAM, Twardzik DR, Vedvick TS;

PI WPI: 1998-261042/23.

DR N-PSDB; AAW64530.

XX Immunogenic Mycobacterium tuberculosis polypeptide(s) and DNA - used  
 PT to develop products for the detection of M. tuberculosis infection  
 PT and for diagnosis, treatment and prevention of tuberculosis

XX Example 3c: Page 182; 230pp: English.

CC This sequence represents an immunogenic portion of a soluble  
 CC Mycobacterium tuberculosis (MT) antigen which can be used in a method  
 CC for inducing protective immunity against tuberculosis (TB). This sequence  
 CC can be formulated into vaccines and/or pharmaceutical compositions for  
 CC immunising against M. tuberculosis infection or may be used for the  
 CC diagnosis of tuberculosis.

SO Sequence 196 AA:

Query Match 90.0%; Score 27; DB 19; Length 196;  
 Best Local Similarity 83.3%; Pred. No. 6.3e+02;

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Oy 1 GSIGGS 6  
 Db 190 gavs9s 195

## RESULT 34

AAW64370  
 ID AAW64370 standard; Protein: 196 AA.

AC AAW64370:  
 AA64370;

DE 09-NOV-1998 (first entry)

DE Mycobacterium tuberculosis antigen XP36.

XX Tuberculosis; infection; diagnosis; antigen; XP36.  
 KW Mycobacterium tuberculosis strain Erdman.  
 DS M09816645-A2.

PN M09816645-A2.

PD 23-APR-1998.

PE 07-OCT-1997; 97MO-US18214.

PR 13-MAR-1997; 97US-0818111.

PR 11-OCT-1996; 96US-0730522.

PA (CORI-) CORIXA CORP.

PI Campos-Neto A, Dillon DC, Houghton R, Lodes MJ;  
 PI Reed SC, Skeiky YAM, Twardzik DR, Vedvick TS;

PI WPI: 1998-251297/22.

DR N-PSDB; AAW44421.

PD 23-APR-1998.

PE Example 3: Page 193-194; 250pp: English.

CC This polypeptide is encoded by Mycobacterium tuberculosis antigen  
 CC XP36 DNA (see AAW44421) that was isolated from a M. tuberculosis  
 CC strain Erdman genomic DNA expression library using sera from  
 CC patients having extrapulmonary tuberculosis. XP36 bears no  
 CC similarity to known sequences. The invention relates to methods  
 CC and compositions for diagnosing tuberculosis. It provides  
 CC polypeptides (see AAW4421-44423) comprising an antigenic portion of a  
 CC M. tuberculosis antigen, as well as DNA sequences encoding such  
 CC polypeptides, recombinant expression vectors and transformed or  
 CC transfected host cells. Also claimed are methods and diagnostic  
 CC kits for detecting M. tuberculosis infection in a patient using  
 CC these polypeptides, antibodies or oligonucleotide probes and  
 CC primers.

SO Sequence 196 AA:

Query Match 90.0%; Score 27; DB 19; Length 196;  
 Best Local Similarity 83.3%; Pred. No. 6.3e+02;

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Oy 1 GSIGGS 6  
 Db 190 gavs9s 195

Query Match 90.0%; Score 27; DB 19; Length 196;  
 Best Local Similarity 83.3%; Pred. No. 6.3e+02;

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

Oy 1 GSIGGS 6  
 Db 190 gavs9s 195

## RESULT 35

AAV39167  
 ID AAV39167 standard; Protein: 196 AA.

AC AAV39167:  
 AAV39167;

DE 05-NOV-1999 (first entry)

DE M. tuberculosis antigen XP36 2nd predicted amino acid sequence.

DE Mycobacterium tuberculosis; M. tuberculosis; antigen; immunogen;  
 DE immunogenicity; immunization; vaccine; infection;  
 DE immune response; skin test.

DE Mycobacterium tuberculosis.





CC polynucleotides may be used for detection of prostate cancer, chromosome  
 CC identification, as chromosome markers, and for numerous other diagnostic  
 CC or research purposes. The prostate cancer antigens may be used to treat  
 CC disorders such as neutral, immune, muscular, reproductive,  
 CC disorders such as pneumonia, cardiovascular, renal, and proliferative  
 CC disorders, wounds, and ulcers, and for use in the treatment of AIDS to  
 CC AAV57303 represent sequences used in the exemplification of the present  
 CC invention.

XX Sequence 233 AA:

SO

Query Match 90.0% Score 27; DB 21; Length 233;  
 Best Local Similarity 83.3%; Pred. No. 7.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGS 6  
 Db 102 gal99s 107

RESULT 38  
 AAY06887  
 ID AAY06887 standard; Protein: 235 AA.  
 XX AAY06887;  
 XX  
 XX 30-JUN-1999 (first entry)  
 DE  
 DE HMHJ20 polypeptide fragment derived from EST.  
 XX  
 XX HMHJ20 polypeptide; p1m family; cancer; autoimmune disease; asthma;  
 XX rheumatoid arthritis; Alzheimer's disease; AIDS; stroke; gene therapy;  
 XX genetic linkage; chromosome localization.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 XX EP911391-A2.  
 XX  
 XX 28-APR-1999.  
 XX  
 XX 19-OCT-1998: 98BP-0308550.  
 PF  
 XX 27-OCT-1997: 97US-0123184.  
 FR 24-OCT-1997: 97US-0063245.  
 XX  
 XX (SNK ) SMITHKLINE BEECHAM CORP.  
 XX  
 XX Albione EF, Kikly KK;  
 XX  
 XX WPI: 1999-246411/21.  
 DR N-PSDB: AAX32973.  
 XX  
 XX New clone, HMHJ20, useful for diagnosing and treating cancer, AIDS  
 XX and autoimmune diseases  
 XX  
 XX Claim 14, Page 15; 21pp: English.

XX The invention relates to a HMHJ20 polypeptide, a member of the p1m  
 CC family of polypeptides. Host cells transfected with a vector comprising  
 CC the HMHJ20 nucleic acid are used for the recombinant expression of the  
 CC polypeptide. The HMHJ20 nucleic acid may be used as probes or primers to  
 CC detect the presence in, or alterations in, the expression of the  
 CC gene, in a patient suspected of, or afflicted with, a disease. The  
 CC diagnosis of or identification of a predisposition to disease may lead to the  
 CC cancer, autoimmune diseases, asthma, rheumatoid arthritis, Alzheimer's  
 CC disease, AIDS and stroke. Similarly measuring the amount of the  
 CC polypeptide in a sample from a patient can be used to diagnose these  
 CC diseases. The polypeptide may be used to identify its agonists or  
 CC antagonists, and to identify membrane-bound or soluble receptors. A  
 CC fusion protein comprising the HMHJ20 polypeptide and an IgG heavy chain may be  
 CC used in therapy, diagnosis and detection. Polynucleotides are used  
 CC for gene therapy; for determining genetic linkage, genetic variability or

CC alterations in gene expression, and may also be used to localize genes  
 CC on chromosomes. The present sequence represents a HMHJ20 fragment  
 CC derived from EST.

XX Sequence 235 AA:

SO

Query Match 90.0% Score 27; DB 20; Length 235;  
 Best Local Similarity 83.3%; Pred. No. 7.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGS 6  
 Db 80 gal99a 85

RESULT 39  
 AAM53007  
 ID AAM53007 standard; Protein: 246 AA.  
 XX AAM53007;  
 XX  
 XX 03-AUG-1998 (first entry)  
 DE  
 DE Mus musculus I- $\alpha$  protein.  
 XX  
 XX I- $\alpha$ ; Inhibitor of MyD family; treatment; diagnosis; myogenesis;  
 XX defects; abnormal development; disease; cleftocranial dysplasia;  
 XX CDD; Rhabdomyosarcoma; muscle tissue cancer.  
 OS  
 XX Mus musculus.  
 XX  
 XX W09808860-A1.  
 XX  
 XX 05-MAR-1998.  
 PD  
 XX 21-AUG-1997: 97MO-US14780.  
 PF  
 XX 27-AUG-1996: 96US-0704931.  
 PR  
 XX (HUTC-) HUTCHINSON CANCER RES CENT FROD.  
 XX  
 XX Chen CMA, Groudiene M, Kraut N, Weintrub H;  
 XX WPI: 1998-179377/16.  
 DR N-PSDB: AAV21882.  
 XX  
 XX Inhibitor of MyD family proteins - useful for, e.g. treatment and  
 XX diagnosis of defects in myogenesis responsible for abnormal  
 XX development  
 XX  
 XX Disclosure: Pages 74-75; 32pp: English.

XX The sequence is that of murine I- $\alpha$  (inhibitor of MyD family)  
 CC protein. Probes from the gene sequence can be used for determining the  
 CC presence of an I- $\alpha$  protein or analogue, or for detecting I- $\alpha$   
 CC against activity in a test substance. The sequence can be used  
 CC to detect, measure and diagnose defects in myogenesis responsible  
 CC for abnormal development in humans and in animals.  
 CC Including humans, these include the human disorders, in children,  
 CC cleftocranial dysplasia (CDD), Rhabdomyosarcomas and other cancers affecting  
 CC muscle tissue in mammals.

XX Sequence 246 AA:

SO

Query Match 90.0% Score 27; DB 19; Length 246;  
 Best Local Similarity 83.3%; Pred. No. 8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGS 6  
 Db 120 gal99s 125

## RESULT 40

AAW53008

AAW53008 standard; Protein; 251 AA.

XX AAW53008:

DT 03-AUG-1998 (first entry)

DE Mus musculus T-mfb protein.

XX T-mfb; Inhibitor of MyoD family; treatment; diagnosis; myogenesis;  
 XX defects; abnormal development; disease; cleidocranial dysplasia;  
 XX CDD; rhabdomyosarcoma; muscle tissue cancer.

XX Mus musculus.

XX MO9808860-A1.

XX 05-MAR-1998.

XX 21-AUG-1997: 97MO-US14780.

XX 27-AUG-1996: 96US-0704931.

XX (HUTC-) HUTCHINSON CANCER RES CENT FRED.

XX Chen CMA, Groudine M, Kraut N, Weintraub H;

XX WPI: 1998-179377/16.

XX N-PSDB; AAW21283.

XX Inhibitor of MyoD family proteins - useful for, e.g. treatment and  
 XX diagnosis of defects in myogenesis responsible for abnormal  
 XX development

XX Disclosure: Pages 77-78; 92pp; English.

XX The sequence is that of murine T-mfb (inhibitor of MyoD family)  
 CC protein. Probes from the gene sequence can be used for determining the  
 CC presence of the gene in a sample of tissue. The sequence can be used  
 CC to predict the presence of the gene in a sample of tissue. The sequence  
 CC can be used to predict the presence of the gene in a sample of tissue.  
 CC to treat, model and diagnose defects in myogenesis responsible  
 CC for abnormal development and disease conditions in mammals,  
 CC including humans. These include the human disease cleidocranial  
 CC dysplasia (CDD), rhabdomyosarcoma and other cancers affecting  
 CC muscle tissue in mammals.

XX Sequence 251 AA;

Query Match 90.08; Score 27; DB 19; Length 251;

Best Local Similarity 83.38; Pred. NO. 8.2e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSECS 6

Db 120 gelsge 125

Search completed: February 4, 2002, 08:03:05  
 Job time: 169 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:01:10 : Search time 12.9 seconds  
(without alignments)

17.053 Million cell updates/sec

Title: US-09-642-660-11

Perfect score: 1 GSI6GS 6

Sequence: 1 GSI6GS 6

Scoring table: gapop 10.0 , gapext 0.5

Searched: 100059 seqs, 36664827 residues

Total number of hits satisfying chosen parameters: 100059

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Minimum Match 1000  
Listing first 100 summaries

Database: SWISPROT\_39.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	100.0	444	1	P35954 bacillus su
2	100.0	444	1	P35954 bacillus su
3	100.0	444	1	P35954 bacillus su
4	100.0	444	1	P35954 bacillus su
5	100.0	444	1	P35954 bacillus su
6	100.0	444	1	P35954 bacillus su
7	100.0	444	1	P35954 bacillus su
8	100.0	444	1	P35954 bacillus su
9	100.0	444	1	P35954 bacillus su
10	100.0	444	1	P35954 bacillus su
11	100.0	444	1	P35954 bacillus su
12	100.0	444	1	P35954 bacillus su
13	100.0	444	1	P35954 bacillus su
14	100.0	444	1	P35954 bacillus su
15	100.0	444	1	P35954 bacillus su
16	100.0	444	1	P35954 bacillus su
17	100.0	444	1	P35954 bacillus su
18	100.0	444	1	P35954 bacillus su
19	100.0	444	1	P35954 bacillus su
20	100.0	444	1	P35954 bacillus su
21	100.0	444	1	P35954 bacillus su
22	100.0	444	1	P35954 bacillus su
23	100.0	444	1	P35954 bacillus su
24	100.0	444	1	P35954 bacillus su
25	100.0	444	1	P35954 bacillus su
26	100.0	444	1	P35954 bacillus su
27	100.0	444	1	P35954 bacillus su
28	100.0	444	1	P35954 bacillus su
29	100.0	444	1	P35954 bacillus su
30	100.0	444	1	P35954 bacillus su
31	100.0	444	1	P35954 bacillus su
32	100.0	444	1	P35954 bacillus su
33	100.0	444	1	P35954 bacillus su

34	90.0	643	1	TAPA_EPRST
35	90.0	673	1	FOX3_HUMAN
36	90.0	726	1	PAT2_VIBAN
37	90.0	773	1	PICR_PABIT
38	90.0	910	1	FOX2_YEAST
39	90.0	910	1	AUX1_HOVIN
40	90.0	972	1	HEBA_HHEBU
41	90.0	972	1	HEBA_HHEBU
42	90.0	985	1	DPOL_HYV1
43	90.0	999	1	HGP4_HAEIN
44	90.0	1046	1	HUUA_HAEIN
45	90.0	1063	1	HGP1_HAEIN
46	90.0	1066	1	HGP1_HAEIN
47	90.0	1067	1	HGP1_HAEIN
48	90.0	1067	1	HGP1_HAEIN
49	90.0	1067	1	HGP1_HAEIN
50	90.0	1067	1	HGP1_HAEIN
51	90.0	1067	1	HGP1_HAEIN
52	90.0	1067	1	HGP1_HAEIN
53	90.0	1067	1	HGP1_HAEIN
54	90.0	1067	1	HGP1_HAEIN
55	90.0	1067	1	HGP1_HAEIN
56	90.0	1067	1	HGP1_HAEIN
57	90.0	1067	1	HGP1_HAEIN
58	90.0	1067	1	HGP1_HAEIN
59	90.0	1067	1	HGP1_HAEIN
60	90.0	1067	1	HGP1_HAEIN
61	90.0	1067	1	HGP1_HAEIN
62	90.0	1067	1	HGP1_HAEIN
63	90.0	1067	1	HGP1_HAEIN
64	90.0	1067	1	HGP1_HAEIN
65	90.0	1067	1	HGP1_HAEIN
66	90.0	1067	1	HGP1_HAEIN
67	90.0	1067	1	HGP1_HAEIN
68	90.0	1067	1	HGP1_HAEIN
69	90.0	1067	1	HGP1_HAEIN
70	90.0	1067	1	HGP1_HAEIN
71	90.0	1067	1	HGP1_HAEIN
72	90.0	1067	1	HGP1_HAEIN
73	90.0	1067	1	HGP1_HAEIN
74	90.0	1067	1	HGP1_HAEIN
75	90.0	1067	1	HGP1_HAEIN
76	90.0	1067	1	HGP1_HAEIN
77	90.0	1067	1	HGP1_HAEIN
78	90.0	1067	1	HGP1_HAEIN
79	90.0	1067	1	HGP1_HAEIN
80	90.0	1067	1	HGP1_HAEIN
81	90.0	1067	1	HGP1_HAEIN
82	90.0	1067	1	HGP1_HAEIN
83	90.0	1067	1	HGP1_HAEIN
84	90.0	1067	1	HGP1_HAEIN
85	90.0	1067	1	HGP1_HAEIN
86	90.0	1067	1	HGP1_HAEIN
87	90.0	1067	1	HGP1_HAEIN
88	90.0	1067	1	HGP1_HAEIN
89	90.0	1067	1	HGP1_HAEIN
90	90.0	1067	1	HGP1_HAEIN
91	90.0	1067	1	HGP1_HAEIN
92	90.0	1067	1	HGP1_HAEIN
93	90.0	1067	1	HGP1_HAEIN
94	90.0	1067	1	HGP1_HAEIN
95	90.0	1067	1	HGP1_HAEIN
96	90.0	1067	1	HGP1_HAEIN
97	90.0	1067	1	HGP1_HAEIN
98	90.0	1067	1	HGP1_HAEIN
99	90.0	1067	1	HGP1_HAEIN
100	90.0	1067	1	HGP1_HAEIN

## ALIGNMENTS

RESULT 1

ID	YBBA_BACSU	STANDARD:	FRG:	444 AA.
AC	53986:			
AD	01-FEB-1995 (Rel. 31, Created)			
AE	01-FEB-1995 (Rel. 31, Last sequence update)			
AF	20-AUG-2001 (Rel. 40, Last annotation update)			
AG	PURATIVE PYS SYSTEM IIBC COMPONENT YBBA (EC 2.7.1.69).			
AN	YBBA OR IPA-160.			
AO	Bacillus subtilis.			
AP	Bacteria; Firmicutes; Bacillus/Clostridium group;			
AQ	Bacillus/Staphylococcus group; Bacillus.			
AR	_LtrAId-123;			
AS	SEQUENCE FROM N.A.			
AT	STRAIN-168:			
AX	MEDLINE-95020537, PubMed-7934828:			
AY	Glaeser P., Kunst F., Arnold M., Coudart M.P., Gonzales W.,			
BA	Hollo M.F., Ionescu M., Lubochinsky B., Mercellino L., Moszer I.,			
BB	Presecan E., Santanu M., Lubochinsky B., Mercellino L., Moszer I.,			
BC	Presecan E., Lubochinsky B., Moszer I.,			
BD	the region from 325 degrees to 333 degrees.			
BE	mol. Microbiol. 10:371-384(1993).			
BF	1. CATALYTIC ACTIVITY: PROTEIN N-PHOSPHOTRANSFERASE + SUGAR -			
BG	PROTEIN HISTIDINE + SUGAR PHOSPHATE.			
BH	1. SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).			
BI	2. SIMILARITY: CONTAINS A PPS EIIB DOMAIN.			
BJ	3. SIMILARITY: CONTAINS A PPS EIIB DOMAIN.			
BK	-----			
BL	This SWISS-PROT entry is copyright. It is produced through a collaboration			
BM	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
BN	the European Bioinformatics Institute. There are no restrictions on its			
BO	use by non-profit institutions as long as its content is in no way			
BP	modified and this statement is not removed. Usage by and for commercial			
BQ	entities requires a license agreement. (See <a href="http://www.isb-ebc.ch/announce/">http://www.isb-ebc.ch/announce/</a>			
BR	or <a href="http://www.ebi.ac.uk/ncbi/blast/">http://www.ebi.ac.uk/ncbi/blast/</a> ).			
BS	EMBL: X73124; CNA51572.1; -			
BT	EMBL: Z59123; CAB1865.1; -			
BU	PIR: S36671; S39671.			
BV	Subtilisin; B010562; YBBA.			
BW	InterPro: IPR003352; PPS_EIIB.			
BX	Protein: 29278; PPS_EIIB, 1-phosphotransferase system; SUGAR transport;			
BY	transferrase; Phosphorylation; Transmembrane; Complete proteome.			
BZ	DOMAIN			
CA	DOMAIN			
CB	DOMAIN			
CC	DOMAIN			
CD	DOMAIN			
CE	DOMAIN			
CF	DOMAIN			
CG	DOMAIN			
CH	DOMAIN			
CI	DOMAIN			
CJ	DOMAIN			
CK	DOMAIN			
CL	DOMAIN			
CM	DOMAIN			
CN	DOMAIN			
CO	DOMAIN			
CP	DOMAIN			
CQ	DOMAIN			
CR	DOMAIN			
CS	DOMAIN			
CT	DOMAIN			
CU	DOMAIN			
CV	DOMAIN			
CW	DOMAIN			
CX	DOMAIN			
CY	DOMAIN			
CA	DOMAIN			
CB	DOMAIN			
CC	DOMAIN			
CD	DOMAIN			
CE	DOMAIN			
CF	DOMAIN			
CG	DOMAIN			
CH	DOMAIN			
CI	DOMAIN			
CJ	DOMAIN			
CK	DOMAIN			
CL	DOMAIN			
CM	DOMAIN			
CN	DOMAIN			
CO	DOMAIN			
CP	DOMAIN			
CQ	DOMAIN			
CR	DOMAIN			
CS	DOMAIN			
CT	DOMAIN			
CU	DOMAIN			
CV	DOMAIN			
CW	DOMAIN			
CX	DOMAIN			
CY	DOMAIN			
CA	DOMAIN			
CB	DOMAIN			
CC	DOMAIN			
CD	DOMAIN			
CE	DOMAIN			
CF	DOMAIN			
CG	DOMAIN			
CH	DOMAIN			
CI	DOMAIN			
CJ	DOMAIN			
CK	DOMAIN			
CL	DOMAIN			
CM	DOMAIN			
CN	DOMAIN			
CO	DOMAIN			
CP	DOMAIN			
CQ	DOMAIN			
CR	DOMAIN			
CS	DOMAIN			
CT	DOMAIN			
CU	DOMAIN			
CV	DOMAIN			
CW	DOMAIN			
CX	DOMAIN			
CY				

```

AC      PA6317;
AD      01-NOV-1995 (Rel. 32, Created)
DT      01-NOV-1995 (Rel. 32, Last sequence update)
DR      SWISSPROT (Release 32)
DE      PMS SYSTEM CELLULOSE-SPECIFIC IIC COMPONENT (ETIC-CEL) [CELLULOSE-
DN      PERMEASE IIC COMPONENT] (PHOSPHOTRANSFERASE ENZYME II, C COMPONENT).
GE      CEL OR LTIC.
OS      Bacillus subtilis.
OC      Bacteria: Firmicutes: Bacilli/Clostridium group:
OQ      Bacillus/staphylococcus group: Bacillus.
OX      K0B1_TaxID=1423;
RC      SEQUENCE FROM N.A.
RP      STRAIN=168;
RT      MEDLINE=97144336; PubMed=8990303;
RA      Medtsoch S., Glaeser P., Krueger S., Hecker M.;
RT      "Identification and characterization of a new beta-glucoside
RT      utilization system in Bacillus subtilis.";
RL      J. Bacteriol. 179(4):96-106(1997).
RN      SEQUENCE FROM N.A.
RS      STRAIN=SGSC 1A1;
RX      MEDLINE=97124196; PubMed=8969509;
RY      Yoshida K., Shindo K., Sano H., Seki S., Fujimura M., Yanai N.,
RA      Miya Y., Fujita Y.;
RT      "Sequencing of a 65 kb region of the Bacillus subtilis genome
RT      containing the iic and cel loci, and creation of a 177 kb contig
RT      library." Microbiol. Comp. 1(2):311-15(1996).
RC      MICROBIAL COPY #123111-15(311966).
RI      FUNCTION: THIS IS A COMPONENT OF THE PHOSHOENOLPYRUVATE-DEPENDENT
RD      SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYRATE ACTIVE
CC      -TRANSPORT SYSTEM. THE IID DOMAINS CONTAIN THE SUGAR BINDING SITE
CC      AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC      PHOSPHORYLATION SITE (THE DONOR IS PHOSHO-HEP); ITA TRANSFERS ITS
CC      PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC      PHOSPHOGLUCOSE.
CC      -- SIMILARITY: COMPAINS A PTS ETIC DOMAIN.
CC      -- SIMILARITY: COMPAINS A PTS ETIC DOMAIN.
CC      --
CC      THIS SMS-S-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL oustration -
CC      the European Bioinformatics Institute. There are no restrictions on its
CC      use by non-profit institutions as long as its contents and for commercial
CC      use it requires a license agreement (See http://www.isb.slb.ch/announce/
CC      or send an email to license@lsb.slb.ch).
CC      -----
CC      EMBL: Z49923; CA990286.1; ?
CC      EMBL: D83026; BA11584.1; ?
CC      EMBL: Z59123; CA11584.1; ?
CC      SUBMITTER: R603363; CEAS ETIC.
CC      DR      PIR: R603363;
CC      DR      Pfam: PF02178; PTS_ETIC_1
KW      Phosphotransferase system; Sugar transport; Transmembrane;
Complete proteome.
FT      TRANSMEM 31          51          POTENTIAL.
FT      TRANSMEM 72          92          POTENTIAL.
FT      TRANSMEM 104         124          POTENTIAL.
FT      TRANSMEM 138         168          POTENTIAL.
FT      TRANSMEM 170         200          POTENTIAL.
FT      TRANSMEM 218         238          POTENTIAL.
FT      TRANSMEM 246         266          POTENTIAL.
FT      TRANSMEM 292         312          POTENTIAL.
FT      TRANSMEM 339         359          POTENTIAL.
FT      TRANSMEM 379         399          POTENTIAL.
FT      TRANSMEM 407         422          POTENTIAL.
FT      TRANSMEM 452 AA: 4853 AA: 396AE5531486C429 CRC64;
SQ      SEQUENCE

```



```

DR EMBL: U07479; AAC23319.1; ALT:SEQ.
DR TIGR: H10661; .
DR InterPro: IPR000531; TonB_boxC.
DR PROSITE: PS00440; TonB-dependent receptor domain.
DR PROSITE: PS01186; TONB_DEPENDENT_REC_2; 1.
KM Outer membrane: Transpor; TonB box; Multigene family; signal;
KW Receptor; Repeat; Complete proteome.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 999 PROBABLE HEMOGLOBIN AND HEMOGLOBIN-
FT REPEAT 26 33 HAPTOGLOBIN BINDING SITE OF 2-P-T-N.
FT REPEAT 30 33 1 X 4 AN TANDEM REPEATS OF Q-P-T-N.
FT REPEAT 34 37 2.
FT REPEAT 38 41 4.
FT REPEAT 42 45 3.
FT REPEAT 46 49 5.
FT REPEAT 50 53 6.
FT REPEAT 54 57 7.
FT SITE 60 70 TONB_BOX.
FT SITE 982 999 TONB_C-TERMINAL_BOX.
SO SEQUENCE 999 AA; 114690 MW; 1A7AB220092BD7D CRC64;

Query Match 100.0%; Score 30; DB 1; Length 999;
Basic Local Similarity 100.0%; Pval No. 1.4e+02;
Matches 6; Conservativity 0; Mismatches 0; Indels 0; Gaps 0

OY 1 CSLEGS 6
Db 190 GSLEGS 195

RESULT 5
ID HGPB_HAEIN STANDARD: PRF: 999 AA.
CD 087285;
DT 20-AUG-2001 (Reel. 40, Created)
DT 20-AUG-2001 (Reel. 40, Last sequence update)
DT 20-AUG-2001 (Reel. 40, Last annotation update)
DE HEMOGLOBIN AND HEMOGLOBIN-HAPTOGLOBIN BINDING PROTEIN B PRECURSOR.
OS HIGH Hemophilus influenzae.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBI_Taxid=727;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=H1689 / Serotype B.
RA MEDLINE=94827137; PubMed=7456772;
RA RFLN=J01000; Genbank=U07479; Still T.L.;
RT "High a gene encoding a second Hemoglobin or the hemoglobin-haptoglobin-binding protein."
RL Infect. Immun. 66:4733-4741(1998).
CC [2]
CC SEQUENCE FROM N.A.
CC STRAIN=Ela / Serotype B.
CC MOTION D.U.T.: See above.
CC HEMOGLOBIN AND HEMOGLOBIN-HAPTOGLOBIN binding proteins In Hemophilus influenzae.
CC Submitted (Apr-2000) to the EMBL/Genbank/DDJJ databases.
CC -1- FUNCTION: ACTS AS A RECEPTOR FOR HEMOGLOBIN OR THE HEMOGLOBIN/HAPTOGLOBIN COMPLEX OF THE HUMAN HOST AND IS REQUIRED FOR HEME UPTAKE.
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.
CC MISCELLANEOUS: THIS PROTEIN IS SUBJECT TO PHASE-VARIABLE EXPRESSION. THIS MECHANISM IS CALLED SLIP-STRAND MISPAIRING. ADDITION OR LOSS OF CCA REPEAT UNITS WOULD CHANGE THE READING FRAME AND RESULT IN INTRODUCTION OF STOP CODONS DOWNSTREAM OF THE REPEAT REGION. THIS MAY BE A MECHANISM OF REGULATION AND A WAY TO AVOID THE IMMUNOLOGICAL RESPONSE OF THE HOST.
CC -1- SIMILARITY: BELONGS TO THE TONB-DEPENDENT RECEPTOR PROTEIN FAMILY;

```

CC		HEMOCLOBIN/HAPTOGLOBIN BINDING PROTEIN SUBFAMILY.
CC		THIS SWISS-PROT entry is copyrighted, it is produced through a collaboration
CC		between the Bioinformatics Institute, there are no restrictions on its use.
CC		use by non-profit institutions as long as its content is in no way distorted,
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (see <a href="http://www.isb.ch/annouce/">http://www.isb.ch/annouce/</a> ) or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
DR	EMBL:	A067030.1 -
DR	EMBL:	AF252266; AF51630.1 -
DR	InterPro:	IPIR000593: TonB box.
DR	Pfam:	PF00593: TonB_box_1.
DR	PROSITE:	PS00440: TONB_DEPENDENT_REC_2; 1.
KW	Outer membrane; Transport; TonB box; Multigene family; Signal;	
NM	Receptor; Repeat.	
FT	SIGNAL	25 999
FT	CHAIN	25 999
FT	DOMAIN	26 57
FT	REPEAT	26 29
FT	REPEAT	26 29
FT	REPEAT	30 33
FT	REPEAT	30 33
FT	REPEAT	34 37
FT	REPEAT	34 37
FT	REPEAT	38 41
FT	REPEAT	38 41
FT	REPEAT	46 49
FT	REPEAT	46 49
FT	REPEAT	50 53
FT	REPEAT	50 53
FT	REPEAT	54 57
FT	REPEAT	54 57
FT	SITE	66 73
FT	SITE	982 999
FT	VARIANT	46 57
FT	VARIANT	46 57
FT	VARIANT	101 101
FT	VARIANT	106 108
FT	VARIANT	121 122
FT	VARIANT	151 154
FT	VARIANT	304 304
FT	VARIANT	550 560
FT	VARIANT	630 633
FT	VARIANT	633 633
FT	VARIANT	680 689
FT	VARIANT	755 755
FT	VARIANT	782 782
SO	SEQUENCE	999 AA: 114435 MW: 58f631FA5D2685B0 CRC64:
Query Match	100.0%	Score 30. DB 1: Length 999:
D	Conservative	100.0% Pctd No. 1(4e+02)
M	Matches	6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY	1 GSLOGS 6	
DB	193 GSLOGS 198	
RESULT	5	
ID	HGBA_HAEIN	STANDARD: PRT: 1013 AA.
AC	OSRYI2:	
DT	20-AUG-2001 (Rel. 40, Created)	
DT	20-AUG-2001 (Rel. 40, Last sequence update)	
DT	20-AUG-2001 (Rel. 40, Last annotation update)	
DE	HEMOGLOBIN BINDING PROTEIN A PRECURSOR.	
NR	HGBA	
CC	Bacteriophage phiX174	
CC	Bacteria; Proteobacteria; gamma subdivision; Pasteurellalesae;	
CC	Haemophilus influenzae	
OX	NCBI_Taxid=727;	
RP	[1]	
RP	SEQUENCE FROM N.A.	
RP	STRAIN-NTHI 1N12;	

RA MEDLINE=20316037; PubMed=10858226;  
 RA Cope L.D., Hrkal Z., Hansen E.J.;  
 CC Detection of phase variation in expression of proteins involved in  
 CC hemoglobin-haptoglobin binding by nonlysine  
 CC Hemophilus influenzae.  
 RL Infect. Immun. 68:4092-4101(2000).  
 CC -1- FUNCTION: ACTS AS A RECEPTOR FOR HEMOGLOBIN OF THE HUMAN HOST AND  
 CC IS REQUIRED FOR HEME UPTAKE.  
 CC -1- SUBCELLULAR LOCATION: OTHER MEMBRANE.  
 CC -1- MISCELLANEOUS: THIS PROTEIN IS SUBJECT TO PHASE-VARIABLE  
 CC REPEAT REGIONS. THE REPEATS VARY IN THE LENGTH OF THE CCA  
 CC ADDITION OR LOSS OF CCA REPEAT UNITS NOT CHANGING STANDARD SEQUENCING.  
 CC FRAME AND RESULT IN INTRODUCTION OF STOP CODONS DOWNSTREAM OF THE  
 CC REPEAT REGION. THIS MAY BE A MECHANISM OF REGULATION AND A  
 CC WAY TO AVOID THE IMMUNOLOGICAL RESPONSE OF THE HOST.  
 CC -1- SIMILARITY: BELONGS TO THE TONB-DEPENDENT RECEPTOR PROTEIN FAMILY;  
 CC HEMOGLOBIN/HAPTOGLOBIN BINDING PROTEIN SUBFAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL AF221059; AABP0176.1; -  
 CC InterPro: IPR000531; TONB\_boxc.  
 DR Pfam: PF00593; TONB\_boxc; 1.  
 DR PROSITE: PS00430; TONB\_DEPENDENT\_REC\_1; FALSE\_NRG.  
 DR PROSITE: PS01156; TONB\_DEPENDENT\_REC\_2; 1.  
 KW Recet, Membrane, Transport, Tonb box; Multigene family; Signal;  
 KW SIGNAL, Repeat, 1.. 24  
 FT CHAIN 25 1013 HEMOGLOBIN BINDING PROTEIN A.  
 FT DOMAIN 26 57 8 x 4 AA TANDUM REPEATS OF Q-P-T-N.  
 FT REPEAT 30 33 2.  
 FT REPEAT 34 37 3.  
 FT REPEAT 38 41 4.  
 FT REPEAT 42 45 5.  
 FT REPEAT 46 49 6.  
 FT REPEAT 50 53 7.  
 FT REPEAT 54 57 8.  
 FT SITE 67 74 TONB BOX.  
 FT SITE 996 1013 TONB C-TERMINAL BOX.  
 SQ SEQUENCE 1013 AA; 116260 MW; 7699643554ED3C1 CRC64;  
 Query Match 100.0%; Score 30; DB 1; Length 1013;  
 Best Local Similarity 100.0%; Pred. No. 1, del-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSGGGS 6  
 Db 194 GSGGGS 199  
 RESULT 7  
 LSHB\_COT3A STANDARD: PRT; 166 AA.  
 AC P45657;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 20-NOV-2001 (Rel. 40, Last annotation update)  
 DE LUTROPIN BETA CHAIN PRECURSOR (LUTEINIZING HORMONE BETA SUBUNIT) (LSH-  
 DE BETA) (LSH-B) (LH-B).  
 OS Colurnix coturnix japonica (Japanese quail).  
 OC Eukaryota; Metazoa; Arthropoda; Chordata; Vertebrata; Euteleostomi;  
 OC Actopterygii; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Columix.

OX NCBI\_Taxid=93934;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=94252550; PubMed=7515015;  
 RA Ando H., Ishii S.;  
 CC Molecular cloning of complementary deoxyribonucleic acids for the  
 CC beta-subunit of gonadotropin hormone alpha-subunit and luteinizing hormone  
 CC beta-subunit precursor molecules of Japanese quail (Colurnix coturnix  
 CC japonica).  
 RL Gen. Comp. Endocrinol. 93:357-366(1994).  
 CC -1- FUNCTION: PROMOTES SPERMATOGENESIS AND OVULATION BY STIMULATING  
 CC THE TESTES AND OVARIES TO SYNTHESIZE STEROIDS.  
 CC -1- SUBUNIT: HETERO DIMER OF A COMMON ALPHA CHAIN AND A UNIQUE BETA  
 CC CHAIN. EACH COMBINE BIOLOGICAL SPECIFICITY TO THYROTROPIN,  
 CC LUTROPIN AND HUMAN CHORIONIC GONADOTROPIN.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOPROTEIN HORMONES BETA CHAIN  
 CC FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: S70834; AAB30867.1; -  
 DR HSRP; P01233; IHRP.  
 DR InterPro: IPR000359; Cys\_kncl.  
 DR PROSITE: PS00430; TONB\_BOXC.  
 DR InterPro: IPR001545; GLYCOPROTEIN.  
 DR Pfam: PF00007; Cys\_kncl; 1.  
 DR PRINTS: PR00438; GRCYSKNOT.  
 DR SMART: SM00668; GHB; 1.  
 DR PROSITE: PS00261; GLYCOPROTEIN\_HORMONE\_BETA\_1; 1.  
 DR PROSITE: PS00689; GLYCOPROTEIN\_HORMONE\_BETA\_2; 1.  
 KW Hormone, Signal, Glycoprotein, Potential.  
 FT SIGNAL 1 21  
 FT CHAIN 22 166 LUTROPIN BETA CHAIN.  
 FT CARBOHYD 60 160 N-LINKED (GLYCAC. . .) (POTENTIAL).  
 FT DISULFID 56 104 BY SIMILARITY.  
 FT DISULFID 70 119 BY SIMILARITY.  
 FT DISULFID 83 135 BY SIMILARITY.  
 FT DISULFID 85 137 BY SIMILARITY.  
 FT DISULFID 140 147 BY SIMILARITY.  
 SQ SEQUENCE 166 AA; 17030 MW; 6B5293BEC3CFAC7 CRC64;  
 Query Match 93.38; Score 28; DB 1; Length 166;  
 Best Local Similarity 83.38; Pred. No. 58;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSGGGS 6  
 Db 46 GSGGGS 51  
 RESULT 8  
 POXN\_DROME STANDARD: PRT; 425 AA.  
 AC P23758;  
 DT 01-NOV-1991 (Rel. 20, Created)  
 DT 01-NOV-1991 (Rel. 25, Last sequence update)  
 DT 20-NOV-2001 (Rel. 40, Last annotation update)  
 DE PAIRED BOX FOX MEDIO PROTEIN (PAIRED BOX NEURONAL PROTEIN).  
 GN FOX-N.  
 OS Drosophila melanogaster (Fruit fly).  
 CC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 CC Ephyroptera; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 CC Ephyroptera; Diptera; Drosophilidae; Drosophila.  
 OC NCBI\_Taxid=7227;  
 RN [1]





DE CRYIA(D) (CRYSTALLINE ENTOMOCIDAL PROTOXIN) (71 KDA CRYSTAL PROTEIN).  
 ON CRY2AD OR CRYIA(D) OR CRY2.  
 OS Bacillus thuringiensis.  
 CC Bacillus thuringiensis; Bacillus/Clostridium group;  
 CC Bacillus/Clostridium group; Bacillus.  
 CX NCBI\_TaxID=1428.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BR30.  
 RA Choi S.-R., Shin B.-S., Park S.-H.;  
 RL Nucleotide sequence of a new Bacillus thuringiensis cry2-type gene \*;  
 CC EMBL/GenBank/DBJ Databases  
 CC FUNCTION: PROMOTES COLONIZOTIC ACTIVITY BY BINDING TO THE MIDGUT  
 CC EPITHELIAL CELLS OF INSECTS  
 CC -1- DEVELOPMENTAL STAGE: THE CRYSTAL PROTEIN IS PRODUCED DURING  
 CC SPORULATION AND IS ACCUMULATED BOTH AS AN INCLUSION AND AS PART  
 CC OF THE SPORE COAT.  
 CC -1- MISCELLANEOUS: TOXIC SEGMENT OF THE PROTEIN IS LOCATED IN THE  
 CC -1- SIMILARITY: BELONGS TO THE DELTA ENDOTOXIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC certified statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 DR EMBL: AF200816; AAF09583.1; \*  
 DR InterPro: IPR001178; Endotoxin.  
 DR Pfam: PF00555; endotoxin; 1.  
 DR NCBI: 50001517; YK1034W.  
 SQ SEQUENCE 633 AA; 70752 MW; 2A89206711B39CB CRC64;  
 QY 1 GSIGS 6  
 DB 349 GSIGS 354  
 RESULT 11  
 ID ATCC SCRPD STANDARD: PRT: 758 AA.  
 AC P36056  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE HYPOHETICAL 87.9 KDA PROTEIN IN PTM-1RXI INTERGENIC REGION PRECURSOR.  
 RN YK1034W OR YK1447.  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetes; Saccharomycetaceae; Saccharomycetes.  
 CX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=588C.  
 RA Purnelle B., Stala J., van Dyck L., Goffeau A.;  
 RL Analysis of an 11.7 kb DNA fragment of chromosome XI reveals a new  
 CC tRNA gene and four new open reading frames including a Leucine zipper  
 CC protein and a homologue to the yeast mitochondrial regulator Atp2 \*;  
 CC yeast 10:125-130(1994).  
 RN [2]  
 RP SEQUENCE OF 1-570 FROM N.A.  
 RC STRAIN=5288C.  
 RA Purnelle B., Stala J., van Dyck L., Goffeau A.;  
 RL MEDLINE=93127731; Pubmed=1481573;  
 RT The sequence of a 12 kb fragment on the left arm of yeast chromosome

RT XI reveals five new open reading frames, including a zinc finger  
 RT protein and a homologue of the UDP-glucose pyrophosphorylase from  
 RT potato \*;  
 RL Yeast 8:977-986(1992).  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC certified statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch))  
 CC -----  
 DR EMBL: X71632; NOT ANNOTATED COS.  
 DR EMBL: X65584; CAA49258.1; \*  
 DR PIR: S37855; S37855.  
 DR SDD: S0001517; YK1034W.  
 DR InterPro: IPR001841; ZnF\_Eng.  
 DR Pfam: PF00099; ZF\_C3HC4\_1.  
 DR NCBI: 50001517; YK1034W.  
 KW Hypothetical protein; Transmembrane; Signal.  
 FT SIGNAL 1 758  
 FT CHAIN 27 758  
 FT TRANSMEM 400 416  
 FT TRANSMEM 440 456  
 FT TRANSMEM 461 477  
 FT TRANSMEM 503 513  
 FT TRANSMEM 607 623  
 FT TRANSMEM 638 654  
 FT TRANSMEM 654 654  
 FT TRANSMEM 654 654  
 SQ SEQUENCE 758 AA; 87879 MW; F29A61E94P96628 CRC64;  
 QY 1 GSIGS 6  
 DB 522 GSIGS 527  
 RESULT 12  
 ID ATCC SCRPD STANDARD: PRT: 1402 AA.  
 AC Q09891.  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 20-MAR-2001 (Rel. 40, Last annotation update)  
 DE HYPOHETICAL 12C  
 GN SPAC24H1.12C  
 OS Schizosaccharomyces pombe (fission yeast).  
 CC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 CC Schizosaccharomycetes; Schizosaccharomycetaceae;  
 CC Schizosaccharomycetes.  
 CX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972.  
 RA Odell C., Churcher C.M., Bartell B.G., Rajandream M.A., Walsh S.V.;  
 RL Smallell (Nov-1995) to the EMBL/GenBank/DBJ Databases  
 CC -1- FUNCTION: THIS MAGNESIUM DEPENDENT ENZYME CATALYZES THE  
 CC HYDROLYSIS OF ATP COMPLETED WITH THE TRANSPORT OF CALCIUM  
 CC -1- CATALYTIC ACTIVITY: ATP + H2O -> ADP + PHOSPHATE.  
 CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.  
 CC -1- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY  
 CC (E1-E2 ATPASES). SUBFAMILY IV.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;  
  
Query Match 83.3%; Score 28; DB 1; Length 1402;  
Best Local Similarity 83.3%; Pred. No. 4; Se-02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GSLOGS 6  
Db 490 GSLOGS 495  
|:|:|:|  
  
RESULT 13  
ID YJ07\_YEAST STANDARD: PRT; 116 AA.  
AC P40365;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 13-OCT-1995 (Rel. 36, Last annotation update)  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;  
  
Query Match 83.3%; Score 28; DB 1; Length 1402;  
Best Local Similarity 83.3%; Pred. No. 4; Se-02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GSLOGS 6  
Db 490 GSLOGS 495  
|:|:|:|  
  
RESULT 13  
ID YJ07\_YEAST STANDARD: PRT; 116 AA.  
AC P40365;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 13-OCT-1995 (Rel. 36, Last annotation update)  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;  
  
Query Match 90.0%; Score 27; DB 1; Length 149;  
Best Local Similarity 83.3%; Pred. No. 84;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GSLOGS 6  
Db 92 GSLOGS 97  
|:|:|:|  
  
RESULT 15  
ID H1S7\_AZOBAR STANDARD: PRT; 207 AA.  
AC P18787;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;  
  
Query Match 90.0%; Score 27; DB 1; Length 149;  
Best Local Similarity 83.3%; Pred. No. 84;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 1 GSLOGS 6  
Db 92 GSLOGS 97  
|:|:|:|  
  
RESULT 15  
ID H1S7\_AZOBAR STANDARD: PRT; 207 AA.  
AC P18787;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DR EMBL: 267577; CAA91777.1; -.  
DR InterPro: IPR001757; E1-E2\_Atpase.  
DR InterPro: IPR001454; Hydrolase.  
DR Pfam: PF00702; Hydrolase; 1.  
DR PROSITE: PS00154; ATPase\_E1\_E2; 1.  
DR Hydrolase; 1706; Membrane; Phosphorylation; Magnesium; ATP-binding.  
FT TRANSMEM 135 135  
FT TRANSMEM 457 477  
FT TRANSMEM 501 521  
FT TRANSMEM 1066 1086  
FT TRANSMEM 1101 1121  
FT TRANSMEM 1151 1171  
FT TRANSMEM 1315 1335  
FT TRANSMEM 1238 1258  
FT TRANSMEM 1260 1280  
FT MOD\_RES 569 569  
FT PHOSPHORYLATION (BY SIMILARITY).  
SO SEQUENCE 1402 AA; 15935 MW; E8A95A759B83020B CRC64;

\*Cloning of histidine genes of *Acetivibrio brasiliense*: organization of the *hbs* gene cluster and nucleotide sequence of the *hbs* gene. *J. Biol. Chem.* 267:22911-22918 (1992).

CC -1- CATALYTIC ACTIVITY: D-AMINO-4-Y-L-GLUTAMATE + H<sub>2</sub>O → L-GLUTAMATE + 3-(IMIDAZOL-4-YL)-2-OXOPROPYL PHOSPHATE. H(2)O.

CC -1- PATHWAY: SEVENTH STEP IN HISTIDINE BIOSYNTHETIC PATHWAY.

CC -1- SUBCELLULAR LOCATION: CYTOSOL (BY SIMILARITY).

CC -1- SIMILARITY: BELONGS TO THE IMIDAZOLECYCLO-PROSPHATE DEHYDRATASE FAMILY.

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC -1- Send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

DR EMBL: X17435; CNA15478.1; -

DR EMBL: X61207; CNA43515.1; -

DR PIR: J60045; J60045.

DR PIR: S16798; S16798.

DR InterPro: IPR000807; IGPD.

DR Pfam: PF00475; IGPD.

DR PROSITE: PS00954; IGPD, BILIVERDINASE 1; 1.

DR PROSITE: PS00955; IGPD, BILIVERDINASE 2; 1.

KM Histidine biosynthesis: Lyase.

SEQUENCE 207 AA; 22655 MW; 1DSECDIAA08BDB CMC64;

Query Match  
Best Local Similarity: 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSIGGS 6  
1:|||||

Db 201 GRLGSS 206

RESULT 16  
STR-STRGR STANDARD; PRT; 281 AA.

AC P09397;

DR 01-MAR-1989 (Rel. 10, Created)

DR 01-MAR-1992 (Rel. 23, Last sequence update)

DR 01-MAR-1992 (Rel. 23, Last annotation update)

DE STREPTOMYCIN BIOSYNTHESIS PROTEIN STRF.

GN STRF.

OS Streptomyces griseus.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.

OX NCBI\_Taxid=1911;

QX NCBI\_Taxid=1911;

RA Sequences from N.A.

RC SPRAIN-W2-3-11.

RX MEDLINE-9137432; Pubmed-1654502;

RA Mansouri K., Piepersberg W.;

RT 'Genetics of streptomycin production in Streptomyces griseus: nucleotide sequence of five genes, strFGH1, including a phosphatase gene'.

RT Mol. Gen. Genet. 228:453-469(1991).

CC -1- FUNCTION: STRF IS INVOLVED IN THE FORMATION OF N-METHYL-L-GLUTAMATE.

CC -1- PATHWAY: STREPTOMYCIN BIOSYNTHESIS.

CC -1- SIMILARITY: LOW, TO THE AMINO-TERMINAL DOMAIN OF GLUCOSE-6-PHOSPHATE ISOMERASE.

CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC -1- Send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

CC entries requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: Y00459; CNA68518.1; -

DR PIR: S17776; S17776.

DR PROSITE: PS00954; IGPD, BILIVERDINASE 1; 1.

SEQUENCE 281 AA; 31726 MW; 82456C5119B6387C CMC64;

Query Match  
Best Local Similarity: 90.0%; Score 27; DB 1; Length 281;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GSIGGS 6  
1:|||||

Db 91 GSIGGS 96

RESULT 17  
H03\_RAT STANDARD; PRT; 290 AA.

AC Q70453;

DR 15-DEC-1998 (Rel. 37, Created)

DR 15-DEC-1998 (Rel. 37, Last sequence update)

DR 30-MAY-2000 (Rel. 39, Last annotation update)

DE HEME OXYGENASE 3 (EC 1.14.99.3) (HO-3).

GN HMOX3.

OS Rattus norvegicus (Rat).

OC Chordata; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI\_Taxid=10116;

RA MEDLINE-9741004; Pubmed-926719;

RT 'Isolation and characterization of the hmo3 gene from the rat brain that encodes hemoprotein heme oxygenase-3.'

RT Eur. J. Biochem. 247:725-732(1997).

CC -1- FUNCTION: HEME OXYGENASE CLEAVES THE HEME RING AT THE ALPHA METHYLENE BRIDGE TO FORM BILIVERDIN. BILIVERDIN IS SUBSEQUENTLY CONVERTED TO BILIRUBIN BY BILIVERDIN REDUCTASE. HEME OXYGENASE 3 COULD BE IMPLICATED IN SOME HEME-DEPENDENT REGULATORY ROLE IN THE CELL.

CC -1- CATALYTIC ACTIVITY: HEME + 3 H<sub>2</sub>O + O<sub>2</sub> → BILIVERDIN + Fe(2+) + CO + 3 A + 3 H<sub>2</sub>O.

CC -1- SUBCELLULAR LOCATION: MICROSOMAL.

CC -1- TISSUE SPECIFICITY: FOUND IN THE SPLEEN, LIVER, THYMUS, PROSTATE, HEART, KIDNEY, BRAIN AND TESTIS.

CC -1- SIMILARITY: BELONGS TO THE HEME OXYGENASE FAMILY.

CC -1- SIMILARITY: CONTAINS 2 HEME REGULATORY MOTIFS (HRM).

CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/>).

CC -1- Send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch).

DR EMBL: AF058787; AAC14142.1; -

DR InterPro: IPR002051; Heme\_oxygenase.

DR Pfam: PF01126; Heme\_oxygenase; 1.

RA Heme; Oxidoreductase; Microsome; Multigene family.

RT 'Heme oxygenase-3 (HMOX3) is a member of the heme oxygenase family.' *J. Biol. Chem.* 273:255-260 (1998).

FT DOMAIN 255-260 HRM 3 (PROBABLE);

SEQUENCE 290 AA; 32592 MW; 6600235C8A185829 CMC64;

Query Match  
Best Local Similarity: 83.3%; Pred. No. 1.6e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



	Matches	:	Conservative	1:	Mismatches	0:	Indels	0:	Gaps
OY	1	SLGCGS	6						
Db	274	CITLGS	279						
RESULT	20								
H02.RAT		STANDARD:	PRT:	315 AA.					
ID	H02.RAT								
NC	P23711..1891	(Rel. 20, Created)							
DT	01-NOV-1991	(Last sequence update)							
DF	30-MAY-2000	(Rel. 39, Last sequence update)							
DE	HEME OXYGENASE 2 (EC 1.14.99.3) (H02).								
GN	HMOX2.								
OS	Rattus norvegicus (Rat.)								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
CC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.								
NCBI_TaxID=10116;									
RP	SEQUENCE FROM N.A.								
RC	TISSUE-Testis;								
RA	MEDLINE=90237051; PubMed=2185251;								
RZ	Rosenberg M.O., Maines M.D.;								
RT	"Isolation, characterization, and expression in Escherichia coli of a								
RF	cDNA encoding rat heme oxygenase-2.";								
RU	(2) Biol. Chem. 265:7501-7506(1990).								
PN	SEQUENCE FROM N.A.								
RC	STRAIN=SPRAGUE-DAWLEY; TISSUE=Liver;								
RA	MEDLINE=94156193; PubMed=8112599;								
RZ	McClourey W.K. Jr., Maines M.D.;								
RT	"The structure, organization and differential expression of the gene								
RF	encoding rat heme oxygenase-2.";								
RU	(2) Biol. Chem. 265:155-161(1994).								
PN	SEQUENCE OF 142-233 FROM N.A., AND PARTIAL SEQUENCE.								
RC	TISSUE-Testis, and Liver;								
RA	MEDLINE=88139412; PubMed=3343248;								
RZ	Cruse I., Maines M.D.;								
RT	"Evidence suggesting that the two forms of heme oxygenase are								
RF	products of different genes."								
RU	J. Biol. Chem. 263:3348-3353(1988).								
CC	-1- HEME BILANDEIN CLEAVES THE HEME RING AT THE ALPHA								
CC	METHOD BY HEME OXYGENASE 2. HEME BILANDEIN IS SUBSEQUENTLY								
CC	CONVERTED TO BILIVERDIN. THE ACTIVITY OF HEME OXYGENASE IS								
CC	PHYSIOLOGICAL CONDITIONS. THE ACTIVITY OF HEME OXYGENASE IS								
CC	HIGHEST IN THE SPLEEN, WHERE SENESECENT ERYTHROCYTES ARE								
CC	SQUASSTRATED AND DESTROYED.								
CC	-1- FUNCTION: HEME OXYGENASE 2 COULD BE IMPLICATED IN THE PRODUCTION								
CC	OF CARBON MONOXIDE IN BRAIN WHERE IT COULD ACT AS A								
CC	NEUROTRANSMITTER.								
CC	-1- CO + 3 A + 3 H(2)O.								
CC	-1- SUBCELLULAR LOCATION: MICROSMAL.								
CC	-1- TISSUE SPECIFICITY: WIDELY DISTRIBUTED IN BODY WITH A HIGH								
CC	CONCENTRATION IN THE BRAIN.								
CC	-1- INDUCTION: HEME OXYGENASE 2 ACTIVITY IS NON-INDUCIBLE.								
CC	-1- SIMILARITY: BELONGS TO THE HEME OXYGENASE FAMILY.								
CC	-1- SIMILARITY: CONTAINS 2 HEME REGULATORY MOTIFS (HRM).								
CC	-----								
CC	This SWIS-S-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL consortium								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	enterprises requires a license agreement (see http://www.isb-sib.ch/announce/								
CC	or send e-mail to: license@isb-sib.ch).								
DR	EMBL; J05405; AAA11340.1 ;								
DR	EMBL; U05013; AAA19130.1 ;								
DR	EMBL; M18918; AAA11347.1 ;								

	DR	PIR: A29922; A29922.
	DR	InterPro: IP002051; Heme_oxygenase.
	DR	Pfam: PF0126; Heme_oxygenase_1.
	DR	PROSITE: PROSITE_HAEMOXYGENASE_1.
	DR	PROSITE: PS00867; HAEMOXYGENASE_1.
	KM	Heme: oxidoreductase [EC 1.11.1.1].
	PT	BINDING: 44 Microsome family,
	PT	DOMAIN: 263 268 PRXIMAL HEME LIGAND (BY SIMILARITY).
	PT	FOLD: 280 285 HRM 2 (POTENTIAL).
	FT	DOMAIN: 142 146 UNRE 2 (POTENTIAL).
	FT	CONFLICT: 230 232 ONSE -> EFNK (IN REF. 3).
	SC	SEQUENCE: 315 AA; 35762 MW; 9814626DIARC CRC64;
	Query Match	90.0% Score 27; DB 1; length 315;
	Best Local Similarity	83.3%; Pred. No. 1.6e+02;
	Matches	5; Conservative 1; Mismatches 0; Indels 0; Gaps 0
OY	1 GSAGS 6	
Db	274 g16063 279	
	11111	
RESULT 21		
ID	ESTL_STRSC	STANDARD: PRT: 345 AA.
AC	P22266; 1891 (rel. 10, Created)	
DC	1 AUG-1991 (rel. 10, last sequence update)	
DT	01-AUG-1991 (rel. 10, last modification update)	
DE	01-NOV-1997 (rel. 35)	
GN	ENZYME PRECURSOR (EC 3.1.1.1).	
OS	ESTL.	
GC	Streptomyces scabies.	
OC	Bacteria; Firmicutes; Actinobacteria; Actinomycetaceae; Streptomycetes.	
CC	Actinomycetaceae; Streptomycineae; Streptomycetaceae; Streptomycetes.	
RN	[Genl:taxid=1950]	
RP	SEQUENCE FROM N.A., AND SEQUENCE OF 40-69.	
RC	MEDLINE:P017224; PubMed-2234271.	
RA	Raymer G., Willard J.M.A., Schotel J.L.:	
RT	"Cloning, sequencing, and regulation of expression of an extracellular esterase gene from the plant pathogen Streptomyces scabies".	
RT	J Biol Chem. 112:7020-7026(1990).	
RX	X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS).	
SA	Mel Y., Schotel J.L., Derewawa U., Swenson L., Patkar S., Darvenda Z.S.:	
SB	"A novel variant of the catalytic triad in the Streptomyces scabies esterase".	
SC	Biol 2:218-223(1995).	
TC	Induction by zinc.	
TI	-! CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-5 IS THE INITIATOR.	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation at the European Bioinformatics Institute. There are no restrictions on its use for non-profit institutions as long as its content is in no way modified or misrepresented. This record may be used freely in commercial entities requires a license to use ( <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> ) or send an email to <a href="mailto:license@isb.sib.ch">license@isb.sib.ch</a> )	
CC	EMBL: M57297; AAA26743.1; -	
CC	EMBL: M57297; AAA26744.1; ALU_INIT.	
CC	PIR: A37843; A37843.	
CC	PIR: A37843; A37843.	
CC	PDB: 1ESD; 15-OCT-95.	
CC	PDB: 1ESD; 15-OCT-95.	
CC	Hydrolase; Serine protease; Signal; 3D-structure.	
TM	SIGNAL 1 39	

FT	CHAIN	40	345		ESTERASE.
FT	DISULFID	73	103		
FT	DISULFID	156	180		
FT	DISULFID	236	294		
SO	SEQUENCE	345 AA;	36670 MW;	237DC28EE185B0E	CHC64;

Query Match	90.08;	Score 27;	DB 1;	Length 345;
Best Local Similarity	83.38;	Pred. No. 1.9e+02;		
Matches	5;	Conservative	1;	Mismatches 0;
				Indels 0;
				Gaps 0

```

OY      1  GSLGGS 6
        |||||:
Db      140  GSLGCN 145

```

	RESULT	22	
ID	NAGA_VIBCH	STANDARD:	PRT: 378 AA.
NC	O32445: O96TA5:	35	(Created)
DR	20-AUG-2001	(Rel. 40)	Last sequence update
DT	20-AUG-2001	(Rel. 40)	Last annotation update
DE	N-ACETYLGLUCOSAMINE-6-PHOSPHATE DEACETYLASE [EC 3.5.1.25] (GLUCNAc 6-P DEACETYLASE).		
DN	NAGA OR VC0994.		
OS	Vibrio cholerae.		
OC	Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.		
ON	[1]	TaxId=606;	
RP	SEQUENCE FROM N.A.		
RX	STRAIN=EL TOR N16961 / SEROTYPE O1;		
RA	MEDLINE=20406833; PubMed=10955301;		
RA	McDonald L., Feltz B.M., Hickey E.S., Peterson J.D., Umeyan L.A., Gill S.R., Nelson K.E., Head T.D., Tateishi H., Richardson D., Ermolaeva M.D., Vanethavech P., Baas S., Qin H., Dragoi I., Sellers P., McDonald L., Uetachapok T., Fleischman R.D., Newman W.C., White O., Salzberg S.L., Smith H.C., Towell R.R., Nekrasov J.J., Venter J.C., Fraser C.M.;		
RA	DNA sequence of both chromosomes of the cholera pathogen Vibrio		
RA	cholerae strain El Tor N16961.		
RL	Nature 406:477-484(2000).		
CC	-1- CATALYTIC ACTIVITY: N-ACETYL-D-GLUCOSAMINE 6-PHOSPHATE + H(2)O		
CC	=> D-GLUCOSAMINE 6-PHOSPHATE + ACETATE.		
CC	-1- PATHWAY: N-ACETYL GLUCOSAMINE UTILIZATION PATHWAY.		
CC	-1- SIMILARITY: BELONGS TO THE NAGA FAMILY.		
CC	-----		
CC	This Swiss-ProT entry is copyright. It is produced through collaboration		
CC	between the University of Geneva and the Institut Pasteur de Paris. There are no restrictions on its		
CC	use by non-profit institutions as long as its content is in no way		
CC	modified and this statement is not removed. Usage by and for commercial		
CC	entities requires a license agreement (see <a href="http://www.isdb-sib.ch/announce/">http://www.isdb-sib.ch/announce/</a> )		
CC	or send an email to <a href="mailto:license@isdb-sib.ch">license@isdb-sib.ch</a> .		
CC	-----		
CM	EMBL: D81830; DAK2834.1;		
DR	20-AUG-2001	(Rel. 40)	Last sequence update
DR	TIGR: VC0994.		
DR	InterPro: IPRO03764; Naga_1		
DR	Pfam: PF02612; Naga_1		
KM	Hydrolase; Carbohydrate metabolism; Complete proteome.		
FM	CONFLICT 79	T -> I (IN REF. 1)	
SO	SEQUENCE 378 AA; 34506344AAFA92A0F CRC64;		

Query Match	90.08;	Score 27;	DB 1;	Length 378;
Best Local Similarity	83.38;	Pred. No. 2.1e+02;		
Matches	5;	Conservative	0;	Indels 0;
		Mismatches	0;	Gaps 0;

QY	1	GSLGGS	6
		:	
Db	301	GTLGGS	306

	RESULT	23	
ID	POLR_TOBAC	STANDARD:	PRT:    396 AA.
DN	01-JUN-1994	(Rel. 29, Creset)	
DT	01-JUN-1994	(Rel. 29, Last sequence update)	
DE	01-FEB-1996	(Rel. 33, Last annotation update)	
DN	POLYGLACTURONASE PRECURSOR [EC 3.2.1.15] (PG) (PECTINASE).		
GZ	PGI.		
OS	Nicotiana glauca (Common tobacco),		
CC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;		
CC	Spermatophytes; Eudicots; Solanales; Nicotianaceae; Nicotiana,		
OC	Species: easterickii L.; Solanales; Solanaceae; Nicotiana,		
OX	[1]		
RN	NBI_LTaxID=1097;		
RF	SEQUENCE FROM N.A.		
RS	STRAIN=C.V. HAVANA; TISSUE=Pollen;		
BX	MEDLINE=14285651; PubMed=6018876;		
TX	"Characterization of a tobacco gene encoding a pollen-specific polygalacturonase";		
PL	Plant Mol. Biol. 25:283-297(1994).		
RL	-1- FUNCTION: MAY FUNCTION IN DEPOLYMERIZING PECTIN DURING POLLEN DEVELOPMENT, GERMINATION, AND TUBE GROWTH.		
CC	-1- CATALYTIC ACTIVITY: RANDOM HYDROLYSIS OF 1,4-ALPHA-D-		
CC	-1- SUBCELLULAR LOCATION: CYTOSOL AND OTHER GALACTUROMANS.		
CC	-1- TISSUE SPECIFICITY: POLLEN.		
CC	-1- SIMILARITY: BELONGS TO FAMILY 28 OF GLYCOSYL HYDROLASES (POLYGALACTUROMANSES).		
CC	-- -- THIS SWISS-PROT entry is copyrighted by its original source(s). All rights reserved. Redistribution and reuse are permitted without further permission provided the European Bioinformatics Institute, where there are no restrictions on ways it can be used by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isdb.ch/commercial or send an email to liceneseis@sb.slb.ch).		
CX	EMBL: X71070; CA503136.1; ALT INIT.		
DR	EMBL: X71017; CA503135.1; ALT INIT.		
DR	EMBL: X71018; CA503136.1; ALT INIT.		
DR	EMBL: X71019; CA503134.1; ALT INIT.		
DR	EMBL: X71016; CA503137.1; ALT INIT.		
PIR	S52007; S52007.		
InterPro	IPE000743; polyalecturonase.		
Drhm	F00230; GlycoLyticEnzyme1.		
KW	Hydrolase; Glycosylase; Cell wall; Signal; glycoprotein.		
FT	SIGNAL	1..22	POTENTIAL.
FT	CHAIN	23..396	POLYGALACTURONASE.
FT	ACT_SITE	236..236	PROBABLE.
FT	CARBOND	143..143	N-LINKED (GLCNAc . ) (POTENTIAL).
FT	CARBOND	151..151	N-LINKED (GLCNAc . ) (POTENTIAL).
FT	CARBOND	164..164	N-LINKED (GLCNAc . ) (POTENTIAL).
FT	CARBOND	184..184	N-LINKED (GLCNAc . ) (POTENTIAL).
FT	CARBOND	203..203	N-LINKED (GLCMAc . ) (POTENTIAL).
FT	CARBOND	208..208	N-LINKED (GLCMAc . ) (POTENTIAL).
FT	CARBOND	259..259	N-LINKED (GLCMAc . ) (POTENTIAL).
FT	CARBOND	294..294	N-LINKED (GLCNAc . ) (POTENTIAL).
FT	VARIANT	154..154	I -> V.
FT	VARIANT	258..258	R -> S.



DB 331 GSWGS 336

RESULT 26  
YCDG\_ECOLI  
ID YCDG\_ECOLI STANDARD: PRT: 404 AA.  
AC P55705;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE HYPOTHEICAL TRANSPORT PROTEIN Y4M.  
OS RIBITOLIDIUM SP. (Strain NGR234).  
OC Plasmid sym pNCR234.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Rhizobium.  
OC NCBI\_TaxID:394;  
RN SEQUENCE FROM N.A.  
RX MEDLINE:9705956; PubMed:9161342;  
RA Freiberg C.A., Folley R., Baltoch A., Broughton W.J., Rosenthal A.,  
RA Perret X.;  
RL Molecular basis of symbiosis between Rhizobium and Legumes.\*;  
CC Nature 387:394-401(1997).  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN. INNER MEMBRANE  
CC (POTENTIAL).  
CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.  
CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.  
CC This SMIS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AE000106; AA091936.1; -;  
DR InterPro: IP800362; sub\_transport.  
DR Pfam: PF00083; suger\_tf.1.  
KW Hypothetical protein; Transport; Transmembrane; Inner membrane;  
KW Plasmid.  
FT TRANSMEM 11 31 POTENTIAL.  
FT TRANSMEM 46 65 POTENTIAL.  
FT TRANSMEM 89 109 POTENTIAL.  
FT TRANSMEM 151 171 POTENTIAL.  
FT TRANSMEM 178 198 POTENTIAL.  
FT TRANSMEM 211 231 POTENTIAL.  
FT TRANSMEM 253 273 POTENTIAL.  
FT TRANSMEM 307 327 POTENTIAL.  
FT TRANSMEM 331 351 POTENTIAL.  
FT TRANSMEM 366 386 POTENTIAL.  
FT TRANSMEM 377 397 POTENTIAL.  
FT TRANSMEM 404 424 POTENTIAL.  
SQ SEQUENCE 404 AA: 42013 MW: C24345EE33730E31 CRC64:

Query Match 90.0% Score 27; DB 1; Length 404;  
Best Local Similarity 83.3% Pred. No. 2-Se-02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
OC NCBI\_TaxID:562;  
RN SEQUENCE FROM N.A.  
RX MEDLINE:97426617; PubMed:9278503;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Colado-Vides J., Glaser J.D., Goeck C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
RN SEQUENCE FROM N.A.  
RP STRAIN-K12;  
RC MEDLINE-97061202; PubMed-8905232;  
RA Oshima T., Aiba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
RA Kikuchi N., Kimura M., Kishino M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
RA Mori H., Motomura K., Nakamura Y., Nishio Y., Saito N.,  
RA Sempel G., Seki Y., Tagami H., Takemoto H., Wada C., Yamamoto Y.,  
RA Yano M., Horikuchi T.;  
RL A 718-kb DNA sequence of the Escherichia coli K-12 genome  
RT corresponding to the 12,728,000 min region on the linkage map.\*;  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (PROBABLE).  
CC -1- SIMILARITY: BELONGS TO THE XANTHINE/URACIL PERMEASES FAMILY.  
CC This SMIS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL: AE000202; AAC74091.1; ALT\_INIT.  
DR EMBL: D90783; BA35773.1; -;  
DR EMBL: D90783; BA35773.1; -;  
DR EMBL: D90783; BA35773.1; -;  
DR EMBL: D90783; BA35773.1; -;  
DR EMBL: D90783; BA35773.1; -;  
DR Pfam: PF00860; xan\_ur\_permease.1.  
KW POSITIVE: P50116; XANTH\_URACIL\_PERMASE; 1.  
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
KW TRANSMEM 58 78 POTENTIAL.  
FT TRANSMEM 80 100 POTENTIAL.  
FT TRANSMEM 111 131 POTENTIAL.  
FT TRANSMEM 151 171 POTENTIAL.  
FT TRANSMEM 170 190 POTENTIAL.  
FT TRANSMEM 197 217 POTENTIAL.  
FT TRANSMEM 241 261 POTENTIAL.  
FT TRANSMEM 285 305 POTENTIAL.  
FT TRANSMEM 319 339 POTENTIAL.  
FT TRANSMEM 348 368 POTENTIAL.  
FT TRANSMEM 386 406 POTENTIAL.  
FT TRANSMEM 407 427 POTENTIAL.  
SQ SEQUENCE 442 AA: 45557 MW: 350C612AE6B0D67 CRC64:

Query Match 90.0% Score 27; DB 1; Length 442;  
Best Local Similarity 83.3% Pred. No. 2-Se-02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;



DR	EMBL	D13382	BAA02550.1	...
DR <th>EMBL</th> <th>AF203791</th> <th>AAAF14001.1</th> <th>...</th>	EMBL	AF203791	AAAF14001.1	...
DR <th>EMBL</th> <th>MS9445</th> <th>AAAF7464.1</th> <th>...</th>	EMBL	MS9445	AAAF7464.1	...
DR <th>EMBL</th> <th>U13873</th> <th>AAAF7463.1</th> <th>...</th>	EMBL	U13873	AAAF7463.1	...
DR <th>EMBL</th> <th>U13873</th> <th>AAAF7463.1</th> <th>...</th>	EMBL	U13873	AAAF7463.1	...
DR <th>PIR</th> <th>A38260</th> <th>A38260.1</th> <th>...</th>	PIR	A38260	A38260.1	...
DR <th>MDJ</th> <th>NCI798257</th> <th>SECV</th> <th>...</th>	MDJ	NCI798257	SECV	...
DR <th>InterPro</th> <th>IPR000087</th> <th>Collagen</th> <th>...</th>	InterPro	IPR000087	Collagen	...
DR <th>InterPro</th> <th>IPR001190</th> <th>SRCR</th> <th>...</th>	InterPro	IPR001190	SRCR	...
DR <th>Pfam</th> <th>PF01391</th> <th>Collagen</th> <th>...</th>	Pfam	PF01391	Collagen	...
DR <th>Pfam</th> <th>PF00350</th> <th>SRCR</th> <th>...</th>	Pfam	PF00350	SRCR	...
DR <th>PRINTS</th> <th>PRO0358</th> <th>SPRACRPTFR</th> <th>...</th>	PRINTS	PRO0358	SPRACRPTFR	...
DR <th>PROSITE</th> <th>PS00420</th> <th>SRCR_1</th> <th>...</th>	PROSITE	PS00420	SRCR_1	...
DR <th>PROSITE</th> <th>PS00287</th> <th>SRCR_2</th> <th>...</th>	PROSITE	PS00287	SRCR_2	...
DR <th>Transmembrane</th> <th>glycoprotein</th> <th>...</th> <th>...</th>	Transmembrane	glycoprotein	...	...
KM <th>Receptor</th> <th>Alternative splicing</th> <th>...</th> <th>...</th>	Receptor	Alternative splicing	...	...
FT <th>DOMAIN</th> <th>1</th> <th>55</th> <th>...</th>	DOMAIN	1	55	...
FT <th>TRANSMEM</th> <th>56</th> <th>78</th> <th>...</th>	TRANSMEM	56	78	...
FT <th>DOMAIN</th> <th>79</th> <th>458</th> <th>...</th>	DOMAIN	79	458	...
FT <th>DOMAIN</th> <th>79</th> <th>458</th> <th>...</th>	DOMAIN	79	458	...
FT <th>DOMAIN</th> <th>209</th> <th>255</th> <th>...</th>	DOMAIN	209	255	...
FT <th>DOMAIN</th> <th>277</th> <th>350</th> <th>...</th>	DOMAIN	277	350	...
FT <th>DOMAIN</th> <th>357</th> <th>457</th> <th>...</th>	DOMAIN	357	457	...
FT <th>DOMAIN</th> <th>357</th> <th>457</th> <th>...</th>	DOMAIN	357	457	...
FT <th>CARBOHYD</th> <th>94</th> <th>94</th> <th>...</th>	CARBOHYD	94	94	...
FT <th>CARBOHYD</th> <th>107</th> <th>107</th> <th>...</th>	CARBOHYD	107	107	...
FT <th>CARBOHYD</th> <th>147</th> <th>147</th> <th>...</th>	CARBOHYD	147	147	...
FT <th>CARBOHYD</th> <th>188</th> <th>188</th> <th>...</th>	CARBOHYD	188	188	...
FT <th>CARBOHYD</th> <th>223</th> <th>223</th> <th>...</th>	CARBOHYD	223	223	...
FT <th>CARBOHYD</th> <th>272</th> <th>351</th> <th>...</th>	CARBOHYD	272	351	...
FT <th>VARSPLIC</th> <th>352</th> <th>354</th> <th>...</th>	VARSPLIC	352	354	...
FT <th>VARSPLIC</th> <th>355</th> <th>458</th> <th>...</th>	VARSPLIC	355	458	...
FT <th>CONFLICT</th> <th>47</th> <th>47</th> <th>...</th>	CONFLICT	47	47	...
FT <th>CONFLICT</th> <th>110</th> <th>110</th> <th>...</th>	CONFLICT	110	110	...
FT <th>CONFLICT</th> <th>120</th> <th>120</th> <th>...</th>	CONFLICT	120	120	...
FT <th>CONFLICT</th> <th>130</th> <th>130</th> <th>...</th>	CONFLICT	130	130	...
FT <th>CONFLICT</th> <th>169</th> <th>169</th> <th>...</th>	CONFLICT	169	169	...
FT <th>CONFLICT</th> <th>169</th> <th>169</th> <th>...</th>	CONFLICT	169	169	...
FT <th>CONFLICT</th> <th>202</th> <th>202</th> <th>...</th>	CONFLICT	202	202	...
SO <th>SEQUENCE</th> <th>458 AA</th> <th>50130 MW</th> <th>...</th>	SEQUENCE	458 AA	50130 MW	...

Query Match

Best local similarity

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0

1 GSIGAS 6

346 GSIGAS 351

90.04; Score 27; DB 1; Length 458;

83.33; Pred. No. 2.6e-02;

RESULT 28

CHLN\_SYNP7 STANDARD; PRT; 466 AA.

AC P14208; 1996 (Ref. 34; Created)

AF 01-09-1998 (Ref. 10; Last sequence update)

DE 20-AUG-2001 (Ref. 10; Last sequence update)

DE 20-AUG-2001 (Ref. 10; Last sequence update)

DE (EC 1.3.1.33) (LI-POR SUBUNIT N) (DOK SUBUNIT N)

GN CHLN.

OS *Synechococcus* sp. (strain PCC 7942) (Anacystis nidulans f2).

OC Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.

NCBI\_TaxID=1140;

SP SOURCE FROM N.A.

RA Lieberman-Hurtwitz J., Rosen-Tarazi M., Gabai C., Hassidim M., Schwartz R., Kaplan A.;

CC Submitted (SEP-1992) to the EMBL/GenBank/DBJ databases.

CC -!- FUNCTION: INVOLVED IN THE LIGHT-INDEPENDENT REDUCTION OF PROTOCHLOROPHYLLIDE (PCHLIDE) TO CHLOROPHYLLIDE (CHLIDE) (BY

PROTEOCHLOROPHYLLIDE + NADPH.  
 CC -1- PATWAY: CHLOROPHYLL BIOSYNTHESIS.  
 CC -1- SIMILARITY: BELONGS TO THE CHLH / DCBH FAMILY.  
 CC -----  
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
 CC BETWEEN THE SWISS INSTITUTE OF BIOFORMATICS AND THE EUROPEAN BIOINFORMATICS  
 CC CENTRE. NON-PROFIT INSTITUTIONS AS LONG AS THE CONTENT IS IN NO WAY  
 CC USED BY NON-PROFIT INSTITUTIONS ARE NOT REMOVED. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X67694; CAA47924.1; *Chlorophyll b* biosynthesis; Oxidoreductase; NADP.  
 CC PROTEINLENGTH: 468 AA; 51940 MW; 981A51305045574C CRC64;  
 CC SQ  
 Query Match 90.0%; Score 27; DB 1; Length 466;  
 Best Local Similarity 83.3%; Pred. No. 2.6e+02; Mismatches 0; Gaps 0;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOC6 6  
 Db 451 GSLOC6 456  
 RESULT 29  
 KZCB\_RAT STANDARD: PRT: 482 AA.  
 ID KZCB\_RAT  
 AC 01-OCT-1996 (Rel. 34, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE KERATIN, TYPE II CYTOSKELETAL 8 (CYTOKERATIN B) (CYTOKERATIN ENDO A).  
 GN KER8.  
 OS Rattus norvegicus (Rat).  
 OC Euteleostomi; Chordata; Craniota; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Scuriognathi; Muridae; Rattus.  
 CX NCBI\_Taxid=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SPRAGE-DANLEY; TISSUE=Prostate;  
 RC MEDLINE=92129304; PubMed=6030145.  
 RC Ando S, Tokui T, Yano T, Inagaki M;  
 RT "Regulation of basal and luminal cell-specific cytokeratin expression  
 RT in rat accessory sex organs. Evidence for a new class of androgen-  
 RT repressed genes and insight into their pairwise control.";  
 RT J. Biol. Chem. 267:2303-2310(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=95042904; PubMed=7525090;  
 RA Wong X, Hsieh J.-T., Zhu H.E.;  
 RT "Cloning and characterization of a specific cytokeratin-8 cDNA from  
 RT rat prostatic epithelium.";  
 RT Chung-Kuo I Hsieh Ko Hsieh Hsueh Pao 16:1-7(1994).  
 RN [3]  
 RP SEQUENCE OF 8-17; 23-74 AND 414-451.  
 RC MEDLINE=95042159; PubMed=8660345.  
 RA Ando S, Tokui T, Yano T, Inagaki M;  
 RT "Keratin 8 phosphorylation in vitro by cAMP-dependent protein kinase  
 RT occurs within the amino- and carboxyl-terminal end domains.";  
 RT Blochem. Biophys. Res. Commun. 221:67-71(1996).  
 CC -1- SUBUNIT: HETEROHEXAMER OF TWO TYPE I AND TWO TYPE II KERATINS.  
 CC -1- KERATIN 8 ASSOCIATES WITH KERATIN 5 OF CYTOSKELETAL AND  
 CC MICROFIBRILLAR KERATIN, I (ACIDIC) AND II (NEUTRAL TO BASIC)  
 CC (40-55 AND 56-70 KILODALTONS, RESPECTIVELY).  
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC -----  
 CC THIS SWISS-PROT ENTRY IS COPYRIGHT. IT IS PRODUCED THROUGH A COLLABORATION  
 CC BETWEEN THE SWISS INSTITUTE OF BIOFORMATICS AND THE EUROPEAN BIOINFORMATICS  
 CC CENTRE. NON-PROFIT INSTITUTIONS AS LONG AS THE CONTENT IS IN NO WAY  
 CC USED BY NON-PROFIT INSTITUTIONS ARE NOT REMOVED. Usage by and for commercial

modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: X63482; AAA19667.1; *ALT-INT*.  
 CC EMBL: X63482; AAA19668.1; *ALT-INT*.  
 CC EMBL: S76054; ?; NOT ANNOTATED\_CDS.  
 CC InterPro: IPRO01664; IF.  
 CC DR InterPro: IPRO03054; Keratin\_11.  
 CC Pfam: PF00038; filament; 1.  
 CC DR PRINTS: PR01226; ?; Colled coil; Keratin; Phosphorylation.  
 CC POSTER: PS00726; ?; Colled coil; Keratin; Phosphorylation.  
 CC INIT MET 0  
 FT DOMAIN 1 89 HEAD.  
 FT DOMAIN 90 397 ROD.  
 FT DOMAIN 398 482 TAIL.  
 FT DOMAIN 90 125 COIL 1A.  
 FT DOMAIN 126 132 LINKER 1.  
 FT DOMAIN 133 158 LINKER 2.  
 FT DOMAIN 159 258 LINKER 12.  
 FT DOMAIN 259 397 COIL 2.  
 FT SITE 341 341 STUTTER.  
 FT MOD\_RES 23 23 PHOSPHORYLATION (BY SIMILARITY).  
 FT MOD\_RES 431 431 PHOSPHORYLATION (BY SIMILARITY).  
 SQ SEQUENCE 482 AA; 53887 MW; C4D0F2FC32C1B5 CRC64;  
 Query Match 90.0%; Score 27; DB 1; Length 482;  
 Best Local Similarity 83.3%; Pred. No. 2.7e+02; Mismatches 0; Gaps 0;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOC6 6  
 Db 446 GSLOC6 451  
 RESULT 30  
 CD5\_HUMAN STANDARD: PRT: 495 AA.  
 ID CD5\_HUMAN  
 AC P06127; *CD5* (Rel. 05, Created)  
 DT 01-JAN-1988 (Rel. 06, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE T-CELL SURFACE GLYCOPROTEIN CD5 PRECURSOR (LYMPHOCYTE GLYCOPROTEIN  
 DE T-CELL-1) (LYMPHOCYTE ANTIGEN CD5).  
 GN CD5 OR LPT0.  
 OS Homo sapiens (human)  
 OS Chordata; Craniota; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CX NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=87014766; PubMed=3093692;  
 RA Jones N.H., Clabdy M.L., Dlayns D.P., Huang H.-J.S.,  
 RT "Isolation of complementary DNA clones encoding the human lymphocyte  
 RT glycoprotein T1/Leu-1.";  
 RT Nature 323:346-349(1986).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lymphocytes; Stimulo M., Vlyas J., Lozano F.;  
 RA Salvo J.;  
 RT "Isolation (FEB 2000) to the EMBL/GenBank/DBD databases.  
 RN [3]  
 RP INTERACTION WITH CD72/LYB-2.  
 RC MEDLINE=9170374; PubMed=171157;  
 RA van de Velde H., van Hoogen T., Luo W., Parnes J.R., Philemans K.;  
 RT "The B-cell surface protein CD72/Lyb-2 is the ligand for CD5.";  
 RT Nature 370:652-655(1992).  
 CC -1- Proliferation. CD5 interacts with CD72/Lyb-2.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN.

CC	-1 SIMILARITY: CONTAINS 3 SRCR DOMAINS.
CC	-1 DATABASE: NAME=PROG; NOTE=cd guide CDS entry;
CC	WWW: <a href="http://www.ncbi.nlm.nih.gov/prov/cds.htm">http://www.ncbi.nlm.nih.gov/prov/cds.htm</a> .
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL contribution "
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/annouces/">http://www.isb-sib.ch/annouces/</a> )
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
DR	EMBL: X04391; CA62797.1; -
DR	EMBL: X89405; CA61584.2; -
DR	EMBL: AJ237937; CA61584.2; JOINED.
DR	EMBL: AJ237928; CA61584.2; JOINED.
DR	EMBL: AJ237929; CA61584.2; JOINED.
DR	EMBL: AJ237930; CA61584.2; JOINED.
DR	EMBL: AJ239311; CA61584.2; JOINED.
DR	EMBL: AJ239312; CA61584.2; JOINED.
DR	PIR: A26396; A26396.1; JOINED.
DR	GlycosylationDB: P06127; -
DR	MIM: 153340; -
DR	InterPro: IPR001190; SRCR.
DR	Pfam: PF00530; SRCR; 3.
DR	SMART: SM00202; SR_2.1; PAUSE_NEG.
DR	PROSITE: PS00287; SRCR.2; 1.
KM	Signal: Transmembrane; Glycoprotein; T-cell; Repeat.
FT	SIGNAL 1 24
FT	CHAIN 25 495
FT	DOMAIN 25 372 T-CELL SURFACE GLYCOPROTEIN CD5.
FT	DOMAIN 373 402 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 403 405 POTENTIAL.
FT	DOMAIN 406 435 POTENTIAL.
FT	DOMAIN 436 458 POTENTIAL.
FT	DOMAIN 459 368 POTENTIAL.
FT	DOMAIN 369 388 SRCR 1.
FT	DOMAIN 389 417 SRCR 2.
FT	DISULFID 44 107 SRCR 3.
FT	CARBOHYD 116 116 N-LINKED (GLCNAC. .) (POTENTIAL).
FT	CARBOHYD 241 241 N-LINKED (GLCNAC. .) (POTENTIAL).
SO	SOURCE 495 AA; 54625 MW; 8740DD76230EBD13 CRGD;

  

QY	1 GSIGS 6	Query Match	50.0%	Score 27.	DP 1:	Length 495;
		Best Local Similarity	83.3%	Prod No. 2	Bes-02;	
Df	179 GSLGGT 184	Matches 5:	Conservative 1;	Mismatches 0;	Indels 0;	Gaps 0;

  

RESULT 31			
HDCAC ARATH	STANDARD:	PRT:	501 AA.
AC OZ2446:			
DT 15-JUN-1998 (Rel. 36, Created)			
DT 15-JUN-1998 (Rel. 36, Last sequence update)			
DI HISTONE H2B (Rel. 18, 10, Last annotation update)			
OS Arabidopsis thaliana (Brassicaceae)			
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC Spermatophytes; Magnoliopsida; eudicotyledons; core eudicots; Rosidae;			
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.			
NCBI_TaxID=3702;			
RN SEQUENCE FROM N.A.			
RA Tomihama T., Shoji K., Hanay H., Okano T.;			
RC STRAIN=CY-COLUMBIA;			
RL Submitted (AUG-1997) to the EMBL/GenBank/DDBJ databases.			
-1 FUNCTION: RESPONSIBLE FOR THE DEACYTYLATION OF LYSINE RESIDUES ON			
THE N-TERMINAL PART OF THE CORE HISTONES (H2A, H2B, H3 AND H4).			
HISTONE DEACETYLTATION PLAYS AN IMPORTANT ROLE IN TRANSCRIPTIONAL			

```
CC CC REGULATION, CELL CYCLE PROGRESSION AND DEVELOPMENTAL EVENTS (BY  
CC SIMILARITY).  
CC -1- SOURCEID=LOCATIONS: NUCLEAR (BY SIMILARITY).  
CC -1- SIMILARITY BELOUGS TO THE HISTONE DEACETYLASE / ACDC / APRA  
CC CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed, usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/about/news/euro  
CC or send an email to license@ebi.ac.uk).  
CC CC  
DR EMBL; AF014824; AAB66486.1; -.  
DR InterPro; IPRO00266; His_deacetylase.  
DR Pfam; PF00850; Hist_deacetyl_1; 1.  
DR PRINTS; PR01711; HISRCRFLSE.  
DR PROSITE; PS01711; HISRCRFLSE.  
DR HMMdomes; HMD0097; Nuclear protein.  
DB SQUENCE; 501 AA; 55037 MW; C50AF562A95B06C2 CRC64;  
  
QY 1 CSIGCS 6  
Db 124 GSVDGS 129  
  
RESULT 32  
ID USP_DROME STANDARD; PRT; 508 AA.  
DT 01-FEB-1991 Ref. 17. Created  
DT 01-FEB-1991 Ref. 17. Last sequence update)  
DT 20-AUG-2001 (Ref. 40, Last annotation update)  
DE ULTRASPIRACLE PROTEIN (XK2C) (CHORION FACTOR 1).  
GN USP OR MZM4 OR CFI OR EG: Z2E5.1 OR CC4380.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pelepodia; Neoptera; Diptera; Brachycera; Muscomorpha;  
NCBI_TaxID=7227;  
OX NCBI_TaxID=7227;  
RN RP SEQUENCE FROM N.A.  
RX TISSUE=Larva;  
RC MEDLINE=90384571; PubMed=2169594;  
RX ORO A.E.; McKewen M.; Evans R.M.;  
FT Relationship between the product of the Drosophila ultraspiracle  
FT gene and retinoid X receptor.  
RL Nature 347:298-301(1990).  
[2]  
RN RP SEQUENCE FROM N.A.  
RX STRAIN=CANTON-S;  
RC MEDLINE=9033404; PubMed=2165589;  
RX Heinrich V.C.; Siller T.J.; Lubahn D.B.; Mecluntyre A.; Gilbert L.I.;  
FT Steroid/hydro hormone receptor superfamily member in Drosophila  
FT mammalian homologues.  
RL Nucleic Acids Res. 18:4143-4148(1990).  
[3]  
RN RP SEQUENCE FROM N.A.  
RX STRAIN=OREGON-R;  
RC MEDLINE=20195011; PubMed=10731137;  
RX Baccelli B.G.; Petraitis C.; Vidali S.; Brun C.; Demillieux J.; Cadton E.;  
RX Drenno S.; Gloux S.; Leclaire V.; Mottier S.; Gallibert F.; Borckova D.;  
RX Minina B.; Rafikov S.C.; Louis C.; Siden-Klimas I.; Bolshakov S.;  
RX Papadogiannakis G.; Spanos L.; Cox S.; Madano E.; de Pablos B.;  
RX Mogelkell J.; Peter A.; Schieffeler P.; Warner M.; Mourikofl F.
```



"A mycoplasma high-affinity transport system and the in vitro invasiveness of mouse sarcoma cells.";  
 RT EMBL J. 73963-3970(1988).  
 CC -1- FUNCTION: PROBABLY PART OF A HIGH-AFFINITY TRANSPORT SYSTEM.  
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC PROTEIN-DEPENDENT TRANSPORT SYSTEMS. CONTAINS TWO DOMAINS.  
 CC THIS SMSIS-PRIOR entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: X14140; CAA3358.1;  
 DR EMBL: M37339; AAB25429.1;  
 DR PIR: S01830; S01830.  
 DR InterPro: PF000313; BRD\_TRANSP.  
 DR PIR: S01830; S01830.  
 KW Transport; Transmembrane; Repeat.  
 FT TRANSMEM 22 42 POTENTIAL.  
 FT TRANSMEM 83 103 POTENTIAL.  
 FT TRANSMEM 120 140 POTENTIAL.  
 FT TRANSMEM 150 170 POTENTIAL.  
 FT TRANSMEM 214 224 POTENTIAL.  
 FT TRANSMEM 234 244 POTENTIAL.  
 FT TRANSMEM 295 315 POTENTIAL.  
 FT TRANSMEM 349 369 POTENTIAL.  
 FT TRANSMEM 403 423 POTENTIAL.  
 FT TRANSMEM 464 484 POTENTIAL.  
 FT TRANSMEM 490 510 POTENTIAL.  
 FT TRANSMEM 527 547 POTENTIAL.  
 FT REPEAT 213 280 APPROXIMATE.  
 FT REPEAT 281 350 APPROXIMATE.  
 SO SEQUENCE 580 AA; 75352C04ABD04749 CRC64.  
 Query Match 90.0%; Score 27; DB 1; Length 580;  
 Best Local Similarity 83.3%; Pred. No. 3; 3e102;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOC5 6  
 Db 495 GSVOGS 500  
 RESULT 34  
 ID TBAF\_EPTST STANDARD: PRT: 643 AA.  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 20-AUG-2001 (Rel. 40, Last annotation update)  
 DE THREAD BIOPOLYMER FILAMENT ALPHA SUBUNIT.  
 OS Epistatulus stoullii (Pacific hagfish).  
 OC Epiplatys: Metazoa; Chordata; Craniota; Hyperotreti; Myxiniiformes;  
 OC Epiplatys; Epiplatinae; Epistatidae.  
 CC MYXINIDAE; Epistatidae; Epistatidae.  
 RN [1] Taxid=7165;  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE=slime gland;  
 RX MEDLINE=9513315; PubMed=7535307;  
 RA Koch E.A., Splitzer R.H., Piltchewalla R.B., Parry D.A.D.;  
 RT A unusual intermediate filament subunit from the cytoskeletal hagfish (Epistatulus stoullii).  
 RT J Cell Sci. 107:3133-3144(1994).  
 CC -1- FUNCTION: RELEASED EXTRACELLULARLY INTO SEAWATER AND PROVIDES PHYSICAL AND BIOLOGICAL DEFENSE AGAINST INVASIVE ORGANISM BY MODULATION OF THE VISCOELASTIC PROPERTIES OF MUCUS.  
 CC -1- SUBUNIT: COILED-COIL HETERO-DIMER OF AN ALPHA AND A GAMMA SUBUNIT.

ASSEMBLE INTO 10 NM FILAMENTS. FORMS A MASSIVE, CONICAL, INTERMEDIATE FILAMENT BIOPOLYMER OF APPROXIMATELY 60 CM.  
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 CC THIS SMSIS-PRIOR entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC EMBL: U11865; AAC59661.1;  
 DR InterPro: IPR001664; IF.  
 DR Pfam: PF00038; filament; 2.  
 DR Intermediate filament; Coiled coil.  
 KW Intermediate filament; Coiled coil.  
 FT DOMAIN 1 151 HEAD.  
 FT DOMAIN 151 210 TAIL.  
 FT DOMAIN 513 530 TAIL.  
 FT DOMAIN 193 227 COIL 1A.  
 FT DOMAIN 228 240 LINKER 1.  
 FT DOMAIN 241 341 COIL 1B.  
 FT DOMAIN 342 362 LINKER 12.  
 FT DOMAIN 363 381 COIL 2A.  
 FT DOMAIN 382 399 LINKER 2.  
 FT DOMAIN 399 417 COIL 2B.  
 FT DOMAIN 453 500 POLY-GLY.  
 SO SEQUENCE 66594 AA; P3145930BDFD819C CRC64.  
 Query Match 90.0%; Score 27; DB 1; Length 643;  
 Best Local Similarity 83.3%; Pred. No. 3; 6e102;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOC5 6  
 Db 68 GSVOGS 73  
 RESULT 35  
 ID F303\_HUMAN STANDARD: PRT: 673 AA.  
 DT 04-03-2001 (Rel. 37, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE FORKHEAD PROTEIN O3A (FORKHEAD IN RABDOMYOSARCOMA-LIKE 1) (AF0321)  
 DE FORKHEAD  
 OS F303A OR FKHL1  
 OS Homo sapiens (human)  
 OC Epiplatys: Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 CC NCBI\_Taxid=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98140118; PubMed=9479491;  
 RA Anderson K.J., Vlastakis C.S., Zeksky S., Cavenee W.K., Arden K.C.;  
 RT A forkhead transcriptional factor subfamily.  
 RT Blood 90:3714-3719(1997).  
 CC -1- FUNCTION: PROBABLY TRANSCRIPTION FACTOR THAT MAY TRIGGERS APOPTOSIS BY INDUCING THE EXPRESSION OF GENES THAT ARE CRITICAL FOR CELL DEATH.  
 CC -1- SUBCELLULAR LOCATION: NUCLEAR.



DT	20-ADENYL (Ref. 40, Last annotation update)
DE	PEROXISOMAL HYDRAE-DENITROGENASE-EPIHEMASE (HDE) (MULTIFUNCTIONAL BETA-OXIDATION PROTEIN) (OMP) [1INDICES: 2-ENOYL-COA HYDRATASE (EC 4.2.1.-); D-3-HYDROXYACVCL COA DEHYDROGENASE (EC 1.1.1.-)].
GN	FOX2 OR YKR0908 OR YKJ108.
LN	Saccharomyces cerevisiae (baker's yeast).
OR	Saccharomycetaceae; Saccharomycetes; Saccharomycetales; Saccharomycetidae; Saccharomycetaceae; Saccharomycetes.
OX	NCBT_TaxId=4932;
RN	SEQUENCE FROM N.A.
RF	[1]
RC	STRAIN=528BC;
PC	MEDLINE=93070612; PubMed=1441752;
EX	Watanabe A., Phillipsen P., Watanabe T., Phillipsen P., 1993, J Biol. Chem. 267:6646-6651(1992).
RT	Cerivastatin, a potent inhibitor of HMG-CoA reductase, reveals nine previously unknown open reading frames.*
RL	Yeast 8:749-759(1992).
RN	SEQUENCE FROM N.A.
RF	[2]
RC	MEDLINE=92202110; PubMed=1551874;
EX	Watanabe A., Phillipsen P., Watanabe T., Phillipsen P., 1993, J Biol. Chem. 267:6646-6651(1992).
CC	- FUNCTION: SECOND TRANSCRIPTIONAL ENZYME ACTING ON THE BETA-OXIDATION PATHWAY FOR FATTY ACIDS. POSSESSING HYDRATASE-DEHYDROGENASE-EPIMERASE ACTIVITIES. CONVERTS TRANS-2-ENOYL-COA VIA D-3-HYDROXYACVCL-COA TO 3-KETONYL-COA.
CC	- SUBUNIT: MONOMER (BY SIMILARITY).
CC	- SUBCELLULAR LOCATION: PEROXISOMAL.
CC	- DOMAIN: CONTAINS TWO SFR DOMAINS.
CC	- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES (SDR) FAMILY.
CC	- TITLE: SWISS-PROT entry is copy righted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <a href="http://www.isb.ch/about/news/cn/">http://www.isb.ch/about/news/cn/</a> ) or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .
CC	EMBL: M66165.1..333778.1 -
CC	EMBL: X65128.1..CAA6243.1 -
CC	EMBL: Z58234.1..CAA82079.1 -
CC	PIR: S25322; S25322.
CC	HSPD: P25529; IAHH.
CC	GCD: S0001717; FOX2.
DR	InterPro: IPRO02199; ADH_short.
DR	InterPro: IPRO02199; ADH_short_C2.
DR	InterPro: IPRO02199; ADH_short_C2.
DR	Pfam: PF00106; adh_short_2.
DR	Pfam: PF01575; Mocc.dehydratase.1.
DR	PRINTS: PR00080; SDRFAMILY.
DR	PROSITE: PS00061; ADH_SHORT_2.
DR	PROSITE: PS00342; MICROBODIES_CTR: 1.
KM	Fatty acid metabolism; Multifunctional enzyme; Oxidoreductase; NMD;
FT	DOMAIN: Isomerase; 210
FT	DOMAIN: Short-chain dehydrogenase like 1; 319
FT	SITE: 898 900 900 900 SHORT-CHAIN DEHYDROGENASE LIKE 1.
FT	FT SITE 13 37 MICROBODY TARGETING SIGNAL (POTENTIAL).
FT	FT ACT_SITE 165 165 NMD (BY SIMILARITY).
FT	FT NP_BIND 326 350 NMD (BY SIMILARITY).
FT	FT ACT_SITE 469 489 BY SIMILARITY.
SQ	SEQUENCE 900 AA; 98703 MW; 66FD0DA3C673768 CRC64;

Query Match      90.0%    Score 27    DB 1    Length 900:

Match             83.1%    Pred. No. 5e+02:

Matches          5;        Conservative    0;        Indels        0;        Gaps        0;-

QY 1 GSIGGS 6  
DB 43 GTLGGs 48

RESULT 39  
ID AUXILIARY STANDARD: PRT: 910 AA.  
AC 027974:  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE AUXILIN.  
OS Bos taurus (Bovine).  
OC Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
NCBI\_TaxID:9913;  
[1]  
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
RC TTSUSE-Brain.  
RX MEDLINE:9520355; PubMed:7703342;  
RA Vith N.G.; Sengul, S.A.; Knorr R., Plessmann U., Weber K.,  
RT "Primary structure of the neuronal clathrin-associated protein  
RT auxilin and its expression in bacteria.";  
RL Eur. J. Biochem. 228:297-304(1995).  
CC -1- FUNCTION: BINDS CLATHRIN HEAVY CHAIN AND PROMOTES ITS ASSEMBLY  
CC -1- INTO REGULAR CAGES.  
CC -1- TISSUE SPECIFICITY: BRAIN.  
CC -1- SUBCELLULAR LOCATION: CYTOSOL.  
CC -1- PTM: TARGET FOR COPI-ASSOCIATED CASEIN KINASE II IN VITRO.  
CC -1- SIMILARITY: CONTAINS 1 TENSIN DOMAIN.  
CC -1- SIMILARITY: CONTAINS 1 J DOMAIN.  
-----  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use, but this statement is as long as its content is in no way  
CC modified and this statement is not removed, usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
-----  
DR EMBL: U09237; AAA79037.1; -  
DR InterPro: IPR001823; DnaJ\_N.  
DR Pfam: PF005087; Trp\_phosphatase.  
DR Pfam: PM00221; DnaJ\_1.  
DR SMART: SM00221; DnaJ\_1.  
DR PROSITE: PS00383; TYR\_PHOSPHATASE\_1; UNKNOWN\_1.  
DR PROSITE: PS00383; TYR\_PHOSPHATASE\_2; UNKNOWN\_1.  
DR PROSITE: PS00636; DnaJ\_1; FALSE\_NEG.  
DR PROSITE: PS0076; DnaJ\_2; 1.  
KM Repeat: Phosphorylation.  
FT 33 34 3 3 X 4 AA APPROXIMATE TANDEM REPEATS.  
FT REPEAT 37 40 2.  
FT REPEAT 41 44 3.  
FT DOMAIN 60 344 TENSIN.  
FT DOMAIN 406 414 SH3-BINDING (POTENTIAL).  
FT DOMAIN 463 757 PRO-RICH.  
FT DOMAIN 528 579 POLY-GLY.  
FT DOMAIN 581 757 J-DOMAIN.  
FT DOMAIN 910 952 U-DOMAIN.  
SQ SEQUENCE 910 AA; 95512 MW; BCL5BDCICF487ED CMC64;

Query Match 90.0%; Score 27; DB 1; Length 910;  
Best Local Similarity 83.3%; Pred. No. 5.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

RESULT 40  
ID HGBA\_HAEDU STANDARD: PRT: 972 AA.  
AC 047952:  
DT 20-AUG-2001 (Rel. 40, Created)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last sequence update)  
DT 20-AUG-2001 (Rel. 40, Last annotation update)  
DE HEMOGLOBIN BINDING PROTEIN PRECURSOR.  
OS Homo sapiens.  
OC Hemophylus ducreyi.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.  
OC NCBI\_TaxID:730;  
OX (1)  
RN 11  
RP SEQUENCE FROM N.A.  
RX MEDLINE:9520355; PubMed:7769598;  
RA Elkins C., Chen C.J., Thomas C.E.;  
RT "Characterization of the hbaA locus encoding a hemoglobin receptor  
RT from Haemophilus ducreyi.";  
RL Infect. Immun. 63:2194-2200(1995).  
CC -1- FUNCTION: ACTS AS A RECEPTOR FOR HEMOGLOBIN OF THE HOST AND IS  
CC REQUIRED FOR HEME UPTAKE.  
CC -1- SUBCELLULAR LOCATION: OUTER MEMBRANE.  
CC -1- HEMOGLOBIN/HAPTOGLOBIN BINDING PROTEIN SUPERFAMILY.  
-----  
CC This SWISS-Prot entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use, but this statement is as long as its content is in no way  
CC modified and this statement is not removed, usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
-----  
DR EMBL: U1281; AAA7311.1; -  
DR InterPro: IPR000531; TonB\_boxC.  
DR Pfam: PF00593; TonB\_boxC\_1.  
DR PROSITE: PS01430; TONB\_DEPENDENT\_REC\_1; FALSE\_NEG.  
DR PROSITE: PS01430; TONB\_DEPENDENT\_REC\_2; FALSE\_NEG.  
KM Outer membrane; transpore; TonB box; Multiligand family; Signal;  
KM Receptor.  
FT SIGNAL 1 22 POTENTIAL.  
FT CHAIN 23 972 HEMOGLOBIN BINDING PROTEIN.  
FT SITE 32 39 TONB BOX.  
FT SITE 955 972 TONB C-TERMINAL BOX.  
SQ SEQUENCE 972 AA; 110937 MW; DEDF059F3575DF CMC64;

Query Match 90.0%; Score 27; DB 1; Length 972;  
Best Local Similarity 83.3%; Pred. No. 5.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSIGGS 6  
DB 156 GTLGGs 161

Search completed: February 4, 2002, 08:01:13  
Job time: 57 sec

QY 1 GSIGGS 6  
DB 680 GTLGGs 685





**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:00:52 ; Search time 23.49 Seconds

(without alignments) 19.457 Million cell updates/sec

Title: **US-09-642-660-11**  
Perfect score: 50  
Sequence: 1 GS/GS 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008  
Listing first 100 summaries

Database: **PIR.68.\***  
1: p1r1.\*  
2: p1r2.\*  
3: p1r3.\*  
4: p1r4.\*

Prod. No. is the number of results predicted by chance for have a  
95% or greater chance of the results being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	107	2	G72668
2	30	100.0	281	2	S44230
3	30	100.0	444	2	S39671
4	30	100.0	452	2	F69651
5	30	100.0	931	2	S66574
6	30	100.0	935	2	T48778
7	30	100.0	1351	2	F75622
8	30	100.0	1346	2	A57376
9	30	100.0	1557	2	T02859
10	30	100.0	1562	2	I51242
11	28	93.3	252	2	T48422
12	28	93.3	267	2	S44225
13	28	93.3	445	2	A81853
14	28	93.3	543	2	T47378
15	28	93.3	562	2	F72608
16	28	93.3	569	2	B84470
17	28	93.3	571	2	C83527
18	28	93.3	572	1	YFEC
19	28	93.3	572	2	C64089
20	28	93.3	572	2	D85504
21	28	93.3	572	2	F69651
22	28	93.3	675	2	G66559
23	28	93.3	729	2	G66559
24	28	93.3	749	2	B84603
25	28	93.3	758	2	S37855
26	28	93.3	851	2	A83484
27	28	93.3	882	2	T01168
28	28	93.3	895	2	T02587

30	28	93.3	940	2	J60291
31	28	93.3	1028	2	E85089
32	28	93.3	1127	2	E85386
33	28	93.3	1146	2	T14270
34	28	93.3	1246	2	T14270
35	28	93.3	1293	2	T14259
36	28	93.3	1402	2	S62557
37	27	90.0	1166	2	S50805
38	27	90.0	149	2	S12192
39	27	90.0	180	2	G70897
40	27	90.0	207	2	D60645
41	27	90.0	211	2	C82641
42	27	90.0	211	2	C82641
43	27	90.0	245	2	P84270
44	27	90.0	267	2	S44228
45	27	90.0	277	2	E96769
46	27	90.0	280	2	A35872
47	27	90.0	281	2	S17776
48	27	90.0	283	2	T36525
49	27	90.0	306	2	D96765
50	27	90.0	310	2	T20229
51	27	90.0	312	2	F70731
52	27	90.0	312	2	F70731
53	27	90.0	315	1	A35199
54	27	90.0	325	2	T36651
55	27	90.0	336	2	E86519
56	27	90.0	337	2	D85504
57	27	90.0	349	2	E86530
58	27	90.0	345	2	A37845
59	27	90.0	349	2	A71212
60	27	90.0	349	2	H72033
61	27	90.0	349	2	E86590
62	27	90.0	350	2	T73538
63	27	90.0	350	2	H86371
64	27	90.0	353	2	G75619
65	27	90.0	365	2	F72033
66	27	90.0	365	2	B81505
67	27	90.0	365	2	C86590
68	27	90.0	368	2	H71600
69	27	90.0	378	2	C05649
70	27	90.0	378	2	E82534
71	27	90.0	378	2	E82534
72	27	90.0	385	2	S46530
73	27	90.0	385	2	S46532
74	27	90.0	394	2	S46532
75	27	90.0	396	2	S46533
76	27	90.0	404	1	S16196
77	27	90.0	458	2	B44407
78	27	90.0	463	2	C54842
79	27	90.0	463	2	C54842
80	27	90.0	465	2	S73875
81	27	90.0	465	1	A36044
82	27	90.0	466	1	A36044
83	27	90.0	495	1	A26396
84	27	90.0	501	2	T05640
85	27	90.0	508	2	S11313
86	27	90.0	508	2	S11313
87	27	90.0	508	2	T83177
88	27	90.0	513	2	T10560
89	27	90.0	580	1	S01830
90	27	90.0	598	2	T17333
91	27	90.0	598	2	S63617
92	27	90.0	621	2	T48187
93	27	90.0	621	2	T48187
94	27	90.0	633	2	S10539
95	27	90.0	713	2	T20090
96	27	90.0	716	2	C01627
97	27	90.0	726	2	B39238
98	27	90.0	737	2	S73792
99	27	90.0	766	2	T05999
100	27	90.0	769	2	T22356

FB19 protein - hum  
probable mutator-1  
probable mutator-1  
Ras-GTPase-activat  
ras-GTPase-activat  
probable calcium-c  
hypothetical prote  
hypothetical prote  
l-madaxoglycyl-  
conserved hypothet  
corrinine carboxym  
stf protein - str  
hypothetical prote  
steroid hormone re  
stf protein - str  
nicotinic dehydroge  
hypothetical prote  
hypothetical prote  
hypothetical prote  
heme oxygenase (de  
probable transcrip  
hypothetical prote  
hypothetical prote  
hypothetical prote  
carboxylesterase (p  
probable oligopept  
hypothetical prote  
hypothetical prote  
macrophage scaveng  
hypothetical prote  
probable hemin ABC  
hypothetical prote  
hypothetical prote  
xilin PB1010W - m  
N-acetylglucosamin  
N-acetylglucosamin  
polyglutacturonase  
polyglutacturonase  
probable lipid car  
polyglutacturonase  
advanced glycosyla  
macrophage scaveng  
hypothetical prote  
probable transpor  
adhesin P1 homolog  
protochondriylid  
T-cell surface gly  
histone deacetylase  
usp protein - Eul  
steroid hormone  
hypothetical prote  
transport system P  
hypothetical prote  
cyma protein - Kle  
hypothetical prote  
intermediate prote  
hypothetical prote  
hypothetical prote  
ferric angulectin  
adhesin P1 precurs  
hypothetical prote  
hypothetical prote



R:Glasner, P.; Lubchinsky, E.; Danonin, A  
Submitted to the EMBL Data Library, July 1995  
A:Description: Bacillus subtilis cel operon.  
A:Accession: U05750  
A:Accession: 557758  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1452 <GAs>  
A:Cross-references: EMBL:Z49992; NID:9895746; PID:CAA90286.1; PID:9895749  
C:Genetics:  
A:Gene: illec  
A:Start codon: CTC  
A:Superfamily: phosphotransferase system enzyme II factor II, phosphoenolpyruvate-depend

Query Match 100.0%; Score 30; DB 2; Length 452;  
Best Local Similarity 100.0%; Pred. No. 1.2e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 232 GSLOGS 237

1 GSLOGS 6  
|||||

RESULT 5  
565574  
transferrin-binding protein I precursor - Actinobacillus pleuropneumoniae  
C:Species: Actinobacillus pleuropneumoniae  
C:Accession: 565574; EMBL:Z49992; NID:9895746; PID:CAA90286.1; PID:9895749  
R:Daban, M.; Medrano, A.; Querol, E.  
Biochem. J. 315, 257-264, 1996  
A:Title: Cloning, sequencing and expression of the transferrin-binding protein I gene fr  
A:Reference number: 565574; MUID:96207589  
A:Accession: 565574  
A:Molecule type: DNA  
A:Residues: 452 <GAs>  
A:Cross-references: EMBL:Z49708; NID:9853708; PID:CAA89810.1; PID:9853709  
C:Genetics:  
A:Superfamily: bacterial pathogen transferrin-binding protein; long-dependent receptor &  
C:Keywords: iron transport; membrane protein; metal binding; receptor  
F:1-22/Domain: signal sequence; status predicted <SIG>  
F:23-201/Product: transferrin-binding protein I #status predicted <MKT>  
F:23-201/Domain: long-dependent receptor amino-terminal homology <NKT>  
F:400-931/Domain: long-dependent receptor carboxyl terminal homology <NKT>

Query Match 100.0%; Score 30; DB 2; Length 931;  
Best Local Similarity 100.0%; Pred. No. 2.4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Db 151 GSLOGS 156

1 GSLOGS 6  
|||||

RESULT 6  
748778  
hypothetical protein J3811.260 [Imported] - Neurospora crassa  
C:Species: Neurospora crassa  
C:Accession: J3811.260; EMBL:Z49992; NID:9895746; PID:CAA90286.1; PID:9895749  
R:Schulte, U.; Altm, V.; Hohnsbeil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Wykatura  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24541  
A:Accession: 748778  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 935 <GAs>  
A:Cross-references: EMBL:AL353820; GSPDB:GN00112; NCSP:J3811.260  
A:Experimental source: cosmid contig J3811; strain 74  
C:Genetics:

```

A:Gene: NCSP:13E11.260
A:Map position: 2

Query Match          100.0%; Score 30; DB 2; Length 935;
Pairwise Similarity 100.0%; Pred.No. 2,de=02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 GSLOGS 6
      |||||
Db      700 GSLOGS 705

RESULT 7
Hypothetical protein - Deinococcus radiodurans (strain R1)
C:Species: Deinococcus radiodurans
C>Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #extL_change 21-Mar-2000
C:Accession: F75622
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Utterback, T.; Zaleski, C.;
Science 286, 1571-1577, 1999
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.
A:Reference number: A75250; MUID:20036896
A:Accession: F75622
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1,936 <MHA>
A:Cross-references: DB:EMBL01825; MIM:96460927; PDB:1AF1261.1; PIR:96460927; TIGR:DR0
C:Genetics:
A:Physical source: strain R1
A:Genetics:
A:Gene: DRB0041
A:Map position: megaplasmid
A:Genome: plasmid
A:Note: plasmid MPl

Query Match          100.0%; Score 30; DB 2; Length 936;
Pairwise Similarity 100.0%; Pred.No. 2,de=02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 GSLOGS 6
      |||||
Db      800 GSLOGS 805

RESULT 8
Hypothetical protein B23111.160 [Imported] - Neurospora crassa
C:Species: Neurospora crassa
C>Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #extL_change 20-Oct-2000
C:Accession: T51904
R:Schulze, O.; Aljan, V.; Uehrase, J.; Brandt, P.; Fittmann, B.; Holland, R.; Nyakatura
C:Title: The Neurospora crassa Sequence Database, August 2000
A:Reference number: Z25858
A:Accession: T51904
A:Molecule type: DNA
A>Status: preliminary
A:Residues: 1,1051 <SCH>
A:Cross-references: EMBL:AL391572; GSFDB:GN00116; NCSP:B23111.160
A:Experimental source: BAC clone B23111; strain OR74A
A:Gene: NCSP:B23111.160
A:Map position: 6

Query Match          100.0%; Score 30; DB 2; Length 1051;
Pairwise Similarity 100.0%; Pred.No. 2,de=02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

      1 GSLOGS 6
      |||||
Db      1 GSLOGS 6

```

Db 828 GSIGGS 833

RESULT 9

Probable regulatory protein 322 - rat

C:Species: Leishmania major

C>Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

C:Date: 08-Feb-1996 #sequence\_revision 00-Feb-1996 #text\_change 08-Feb-1996

C:Accession: A57376

A:Accession: I51242

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 116

A:Cross-references: GB:570834; NID:9546921; PDB:1AB30867.1; PDB:9546922

C:Superfamily: pituitary glycoprotein hormone beta chain

Query Match

Best Local Similarity 93.3%; Score 28; DB 2; Length 166;

Matches 5; Conservative 1; Mismatches 0; Gaps 0;

1 GSIGGS 6

Db 46 GSNIGGS 51

RESULT 12

hypothetical protein P8F6\_160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse ear cress)

C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000

C:Accession: T86422

RESULT 14  
 A:Accession: A8153  
 C:Species: Drosophila melanogaster  
 C>Date: 04-Mar-1993 #sequence-revision 18-Nov-1994 #text-change 21-Jul-2000  
 C:Accession: A8153; S06951  
 R:Damjanij-Claidiere, C.; Jamet, E.; Burrl, M.; Bopp, D.; Basler, K.; Hafen, E.; Dumont, N.  
 Cell 69, 159-172, 1992  
 A:Title: The paired box gene *pox* neuro: a determinant of poly-innervated sense organs in  
 A:Reference number: A8153; MUID:9220941  
 A:Molecule type: DNA  
 A:Residues: 1-425 <DM>  
 A:Cross-references: GB:M6927; NID:915815; PID:9158156  
 R:Bopp, D.; Jamet, E.; Baumgartner, S.; Burrl, M.; Noll, M.  
 A:Title: 81941477, 1999 tissue-specific Drosophila paired box genes, *Pox* meso and *Pox*  
 A:Reference number: S06950; MUID:9005940  
 A:Accession: S06951  
 A:Molecule type: DNA  
 A:Residues: 1-135 <RBP>  
 A:Cross-references: EMBL:X58917; NID:98360; PID:CA41721.1; PID:91405550  
 C:Genetics: *Aradidopsis*; *Pox*-*n*  
 A:Cross-references: PLYBase:PM0003110  
 C:Superfamily: paired box transcription factor Pax-B, paired box homology  
 P:5-10/Domain: paired box homology <RBP>

Query Match 93.3%; Score 28; DB 2; Length 425;  
 Best Local Similarity 83.3%; Pred. No. 2, 8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 GSIUGS 6  
 11:1111  
 70 GSIGGS 75

RESULT 15  
 T00450  
 probable monosaccharide transport protein T14N5.7 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 01-Feb-1999 #sequence-revision 01-Feb-1999 #text-change 24-Sep-1999  
 C:Accession: T00450  
 R:Reidspiegel, N.A.; Palm, C.J.; Conway, A.B.; Kurtz, D.B.; Conway, A.R.; Au, M.; Arcajo,  
 A.; et al. J. Biol. Chem. 274:11553-11557, 1999  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: 214152  
 A:Accession: T00450  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-504 <FED>  
 A:Cross-references: EMBL:AC004260; NID:9317694; PID:MAC34349.1; PID:93540199  
 C:Genetics: *Arabidopsis*; *Arabidopsis*  
 A:Map position: 1  
 A:Notes: 48/1; 155/3; 365/3  
 A:Note: T14N5.7  
 C:Superfamily: glucose transport protein  
 C:Keywords: sugar transport; transmembrane protein

Query Match 93.3%; Score 28; DB 2; Length 504;  
 Best Local Similarity 83.3%; Pred. No. 3, 4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 GSIUGS 6  
 11:1111  
 33 GSIGGS 38

RESULT 16  
 E72608

Probable Hydn APE131 - Aeropyrum pernix (strain K1)  
 C:Species: Aeropyrum pernix  
 C>Date: 20-Aug-1999 #sequence-revision 20-Aug-1999 #text-change 20-Jun-2000  
 C:Accession: E72608  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-uo, K.; Tsu-  
 ma, S.; Takahashi, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.  
 A:Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, *Aero*  
 A:Reference number: A72450; MUID:9910339  
 A:Accession: E72608  
 A:Status: Preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-562 <KAW>  
 A:Cross-references: EMBL:AF000061; NID:95104821; PID:BA480323.1; PID:95105009  
 A:Experimental source: Strain K1  
 C:Genetics:  
 A:Gene: APE131  
 C:Superfamily: N-methylhydantoinase (ATP-hydrolyzing)

Query Match 93.3%; Score 28; DB 2; Length 562;  
 Best Local Similarity 83.3%; Pred. No. 3, 8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 GSIUGS 6  
 11:1111  
 136 GSIGGS 141

RESULT 17  
 B84470  
 Mutator-like transposase [Imported] - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C>Date: 02-Feb-2001 #sequence-revision 02-Feb-2001 #text-change 02-Feb-2001  
 C:Accession: B84470  
 R:Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Hatakeyama, Y.; Jin-uo, K.; Tsu-  
 ma, S.; Takahashi, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudo, Y.; Yamazaki, J.  
 A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
 A:Reference number: A84420; MUID:20083487  
 A:Accession: B84470  
 A:Status: Translated  
 A:Molecule type: DNA  
 A:Residues: 1-569 <STO>  
 A:Cross-references: GB:AB002093; NID:94581167; PID:AMD24650.1; GSPDB:GM00139  
 C:Genetics:  
 A:Gene: AC2905570  
 A:Map position: 2

Query Match 93.3%; Score 28; DB 2; Length 569;  
 Best Local Similarity 83.3%; Pred. No. 3, 8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 1 GSIUGS 6  
 11:1111  
 165 GSIGGS 170

RESULT 18  
 C83527  
 PROLYL-tRNA synthetase PA0956 [Imported] - Pseudomonas aeruginosa (strain PA01)  
 C:Species: Pseudomonas aeruginosa  
 C:Accession: C83527  
 R:Stover, C.K.; Pham, X.Q.; Ervin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickley, M.J.;  
 Adman, S.; Yuen, Y.; Brody, L.L.; Conliffe, S.N.; Folger, K.R.; Kass, A.; Lathig, K.; L.  
 ; Lory, S.; Olson, M.V.  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa  
 A:Reference number: A82950; MUID:2043737  
 A:Accession: C83527

A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-571 <GTO>  
 C:Species: Escherichia coli  
 C>Date: 18-Aug-1995 #sequence,revision 18-Aug-1995 #text,change 16-Jul-1999  
 A:Cross-references: GB:AE004529; GB:AE004091; NID:g9946851; PIDD:AA004345.1; GSPDB:GN001  
 A:Experimental source: strain P01  
 C:Genetics: pros  
 C:Superfamily: proline--trna ligase

Query Match 93.3%; Score 28; DB 2; Length 571;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSIGS 6  
 11:1111  
 Db 200 GSIGS 205

RESULT 19  
 TYPE: proline--trna ligase (EC 6.1.1.15) - Escherichia coli  
 M:Alternate names: globl RNA synthetase factor; prolyl--trna synthetase  
 C:Species: Escherichia coli  
 C>Date: 30-Sep-1992 #sequence,revision 31-Oct-1997 #text,change 24-Sep-1999  
 A:Accession: B64744; J00110; J00061  
 R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Berna, N.T.; Burland, V.; Riley, M.; Co  
 S:Rosen, D.; Stahl, B.; Zhao, Y.  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; M01D:97446517  
 A:Accession: B64744  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-572 <GATX>  
 A:Experimental source: strain K-12; substrain M0155  
 R:Erstl, G.; Delaune, M.; Poch, O.; Gangloff, J.; Moras, D.  
 C:Accession: J00110  
 A:Title: Partition of RNA synthetases into two classes based on mutually exclusive sets  
 A:Reference number: J00110; M01D:90370122  
 A:Accession: J00110  
 A:Residues: 1-572, 'TV', 28-204, 'Q', 206-572 <EMD>  
 A:Cross-references: GB:V9758; NID:g44361; PIDD:AAA24420.1; PID:g147362  
 A:Note: the sequence from Fig. 2 is inconsistent with that from Fig. 1 in having 205-see  
 J. Bacteriol. 172, 281-286, 1990  
 R:Zhao, Z.; Szyman, M.  
 A:Title: Identification and sequence of the drpa gene from Escherichia coli.  
 A:Reference number: J00051; M01D:90094229  
 A:Accession: J00051  
 A:Molecule type: DNA  
 A:Residues: 1-126, 'HV', 129-204, 'RPLTNSQRR', 217-515, 'RA', <ZHO>  
 A:Cross-references: GB:M32357; NID:g445802; PIDD:AAA23710.1; PID:g145803  
 A:Note: the authors identified this protein as global RNA synthesis factor (encoded by c  
 C:Genetics: pros  
 C:Superfamily: proline--trna ligase  
 C:Keywords: aminocyl--trna synthetase; AMP; homodimer; ligase; protein biosynthesis

Query Match 93.3%; Score 28; DB 1; Length 572;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSIGS 6  
 11:1111  
 Db 200 GSIGS 205

RESULT 20  
 C64083

Proline--trna ligase (EC 6.1.1.15) - Haemophilus influenzae (strain Rd KW20)  
 M:Alternate names: prolyl--trna synthetase  
 C:Species: Haemophilus influenzae  
 C>Date: 18-Aug-1995 #sequence,revision 18-Aug-1995 #text,change 16-Jul-1999  
 A:Cross-references: GB:U03756; GB:U42023; NID:g1573729; PIDD:AA023388.1; PID:g1573733  
 A:Experimental source: strain Q157:H7  
 C:Genetics: pros  
 C:Superfamily: proline--trna ligase  
 C:Keywords: aminocyl--trna synthetase; ligase; protein biosynthesis

Query Match 93.3%; Score 28; DB 2; Length 572;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSIGS 6  
 11:1111  
 Db 200 GSIGS 205

RESULT 21  
 TYPE: trna synthetase [imported] - Escherichia coli (strain Q157:H7)  
 C:Species: Escherichia coli  
 C>Date: 16-Feb-2001 #sequence,revision 16-Feb-2001 #text,change 31-Mar-2001  
 A:Accession: D85504  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 Litter, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanta, E.; Poulosis, K.; Apoda  
 Nature 409, 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
 A:Reference number: A85480; M01D:21074933; PMID:11200651  
 A:Accession: D85504  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-572 <GTO>  
 A:Cross-references: GB:AE005174; NID:g12512935; PIDD:AA054496.1; GSPDB:GN00145; UNKPF  
 A:Experimental source: strain Q157:H7; substrain ED593  
 A:Genetics: pros  
 C:Superfamily: proline--trna ligase

Query Match 93.3%; Score 28; DB 2; Length 572;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSIGS 6  
 11:1111  
 Db 200 GSIGS 205

RESULT 22  
 TYPE: trna synthetase [imported] - fruit fly (Drosophila melanogaster)  
 C:Species: Drosophila melanogaster  
 C>Date: 11-Jun-1999 #sequence,revision 11-Jun-1999 #text,change 17-Nov-2000  
 A:Accession: T08600  
 R:Mugner, R.A.; Friedman, T.B.; Smith, B.; Wallpeth, L.L.; Johnson, S.; Burnette, J.B.  
 Submitted to the EMBL data library, August 1997  
 A:Reference number: T08600  
 A:Accession: T08600



A>Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Accession: U01807.1; NID:92394383; PID:92394384  
 A:Cross-references: strain Oregon R  
 A:Experimental source: strain Oregon R  
 A:Gene: Opa  
 A:Cross-references: FlyBase:FBgn020880  
 A:Map position: 2

Query Match 93.3%; Score 28; DB 2; Length 654;  
 Best Local Similarity 83.3%; Pred. No. 4.4e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GS/GCS 6  
 DB 387 GS/GCS 382

## RESULT 23

173738  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Accession: U01807.1; NID:92394383; PID:92394384  
 A:Cross-references: strain Oregon R  
 A:Experimental source: strain Oregon R  
 A:Gene: Opa  
 A:Cross-references: FlyBase:FBgn020880  
 A:Map position: 2

Query Match 93.3%; Score 28; DB 2; Length 675;  
 Best Local Similarity 83.3%; Pred. No. 4.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GS/GCS 6  
 DB 180 GS/GCS 185

## RESULT 24

956559  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Accession: U01807.1; NID:92394383; PID:92394384  
 A:Cross-references: strain Oregon R  
 A:Experimental source: strain Oregon R  
 A:Gene: Opa  
 A:Cross-references: FlyBase:FBgn020880  
 A:Map position: 2

A:Gene: FSP19.7  
 A:Map position: 1

Query Match 93.3%; Score 28; DB 2; Length 729;  
 Best Local Similarity 83.3%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GS/GCS 6  
 DB 163 GS/GCS 168

## RESULT 25

886403  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Accession: U01807.1; NID:92394383; PID:92394384  
 A:Cross-references: strain Oregon R  
 A:Experimental source: strain Oregon R  
 A:Gene: Opa  
 A:Cross-references: FlyBase:FBgn020880  
 A:Map position: 2

Query Match 93.3%; Score 28; DB 2; Length 749;  
 Best Local Similarity 83.3%; Pred. No. 5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GS/GCS 6  
 DB 180 GS/GCS 185

## RESULT 26

956559  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Accession: U01807.1; NID:92394383; PID:92394384  
 A:Cross-references: strain Oregon R  
 A:Experimental source: strain Oregon R  
 A:Gene: Opa  
 A:Cross-references: FlyBase:FBgn020880  
 A:Map position: 2





ras GTPase-activating protein, synaptic - rat  
 A:Accession: U01067  
 C:Species: Rattus norvegicus (Norway rat)  
 C>Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 20-Sep-1999  
 R:Accession: T14259  
 R:Chen, H.J.; Rojas-Soto, M.; Kennedy, M.B.  
 submitted to the EMBL data library, February 1998  
 A:Description: A synthetic ras GTPase-activating protein (p135 synGAP) inhibited by Cxk  
 A:Accession: T14259  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Accession: 1-1293 <CH2>  
 A:Cross-references: EMBL:AF048976; NID:q2935447; PID:q2935448; PIDN:MAC08071.1  
 A:Experimental source: strain Sprague-Dawley

Query Match 93.3%; Score 28; DB 2; Length 1293;  
 Best Local Similarity 83.3%; Pred. No. 8.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 1098 GSLOGS 1103

RESULT 36  
 Probable calcium-transporting ATPase - fission yeast (Schizosaccharomyces pombe)  
 A:Accession: U01067  
 C:Date: 16-May-1996 #sequence\_revision 13-Mar-1997 #text\_change 21-Jan-2000  
 R:Odell, C.; Churcher, C.M.; Barrett, B.G.; Rajandream, M.A.; Walsh, S.V.  
 submitted to the EMBL data library, November 1995  
 A:Reference number: 221786  
 A:Accession: T38329  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Accession: 1-1402 <CO2>  
 A:Residues: 1-1402  
 A:Cross-references: EMBL:Z67757; NID:q1061288; PIDN:CA931777.1; PID:q1061300; CSPDN:GN00  
 A:Experimental source: strain 972h; cosmid c24B11  
 C:Genetics:  
 A:Gene: SPAC24B11.12c  
 A:Map position: 1L  
 A:Insertions: 124/3

Query Match 93.3%; Score 28; DB 2; Length 1402;  
 Best Local Similarity 83.3%; Pred. No. 9.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 490 GSLOGS 495

RESULT 37  
 Hypothetical protein Y1067v - yeast (Saccharomyces cerevisiae)  
 A:Accession: U01067  
 A:Name: hypothetical protein Y1067v  
 C:Species: Saccharomyces cerevisiae  
 C>Date: 11-Jan-1995 #sequence\_revision 08-Sep-1995 #text\_change 29-Oct-1999  
 R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.  
 Yeast 11, 57-60, 1995  
 A:Title: Sequence of a 17.1 kb DNA fragment from chromosome X of Saccharomyces cerevisiae  
 A:Accession: U01067  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Residues: 550750; MUID:95282514  
 A:Molecule type: DNA  
 A:Accession: 1-116 <YAN>  
 A:Cross-references: EMBL:Z34288; NID:q498992; PIDN:CA84056.1; PID:q499000  
 A:Experimental source: strain Y1067v; submitted to the EMBL data library, June 1994

R:Vandenbol, M.; Durand, P.; Dion, C.; Portetelle, D.; Hilger, F.  
 submitted to the EMBL data library, June 1994  
 A:Description: Sequence analysis of a 17.1 kb DNA fragment from chromosome X of Sacch  
 A:Reference number: S47117  
 A:Accession: S47124  
 A:Molecule type: DNA  
 A:Accession: 1-116 <YAN>  
 A:Cross-references: EMBL:Z34288; NID:q498992; PIDN:CA84056.1; PID:q499000  
 R:Vandenbol, M.; Durand, P.; Portetelle, D.; Hilger, F.  
 submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S56835  
 A:Accession: S56841  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Accession: 1-1293 <CH2>  
 A:Cross-references: EMBL:Z49342; NID:q1008215; PIDN:CA89357.1; PID:q1008216; MIPS:Y3  
 C:Genetics:  
 A:Map position: 10L  
 C:Superfamily: Saccharomyces hypothetical protein Y1067v

Query Match 90.0%; Score 27; DB 2; Length 116;  
 Best Local Similarity 83.3%; Pred. No. 1.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 27 GSLOGS 32

RESULT 38  
 Hypothetical protein 5 - Thibacillus ferrooxidans plasmid pRF1  
 A:Accession: U01067  
 C:Species: Thibacillus ferrooxidans  
 C>Date: 19-Mar-1997 #sequence\_revision 25-Apr-1997 #text\_change 25-Apr-1997  
 R:Boettger, M.; Sanger, P.; Lau, P.C.K.  
 Mol. Microbiol. 4, 1381-1391, 1990  
 A:Title: The mobilization and origin of transfer regions of a Thibacillus ferrooxidans  
 A:Reference number: S12188; MUID:91125140  
 A:Accession: S12192  
 A:Status: preliminary; translation not shown  
 A:Cross-references: EMBL:Z52699  
 A:Molecule type: DNA  
 A:Accession: 1-149 <DBO>  
 A:Residues: 1-149  
 A:Cross-references: EMBL:Z52699  
 C:Genetics:  
 A:Genome: Plasmid pRF1

Query Match 90.0%; Score 27; DB 2; Length 149;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 92 GSLOGS 97

RESULT 39  
 Hypothetical protein Rv1105 - Mycobacterium tuberculosis (strain H37Rv)  
 A:Accession: U01067  
 C:Species: Mycobacterium tuberculosis  
 C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
 R:Cole, S.T.; Broese, R.; Parhill, J.; Garsner, T.; Churcher, C.; Harris, D.; Gordon  
 Rajandream, M.A.; Rogers, J.; Davillo, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,  
 Nature 393, 537-544, 1998  
 A:Authors: Sagers, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrett, B.C.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: G70857  
 A:Status: preliminary; translation not shown  
 A:Molecule type: DNA  
 A:Accession: 1-116 <YAN>  
 A:Cross-references: EMBL:Z34288; NID:q498992; PIDN:CA84056.1; PID:q499000  
 A:Experimental source: strain Y1067v; submitted to the EMBL data library, June 1994

A:Residues: 1-180 <COL>  
 A:Cross-references: GB:AL021897; GB:AL123456; NID:93256022; PIDD:CAA17221.1; PID:e125196  
 A:Experimental source: strain H37Rv  
 C:Accession: J050045  
 A:Gene: RV1105

Query Match 90.0%; Score 27; DB 2; Length 180;  
 Best Local Similarity 83.3%; Pred. No. 1.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSI/GCS 6  
 |:|||||  
 DB 98 GSI/GCS 103

## RESULT 40

J050045  
 A:Accession: J050045  
 A:Gene: imidazoleglycerol-phosphate dehydratase (EC 4.2.1.19) - Azospirillum brasilense  
 C:Species: Azospirillum  
 C:Date: 07-Jun-1990 #sequence revision 07-Jun-1990 #text-change 31-Mar-2000  
 C:Accession: J050045; S04399; S16798  
 R:Paul, R.; Bazzicalupo, M.; Damiani, G.; Bianchi, A.; Schipani, C.; Sgoramella, V.; Pol  
 M.; Gen. Genet. 216, 224-229, 1989  
 R:title: Cloning of histidine genes of Azospirillum brasilense: organization of the ABH  
 A:Accession number: J050045; NCID:89313660  
 A:Gene: imidazoleglycerol-phosphate dehydratase  
 A:Molecule type: DNA  
 A:Residues: 1-207 <FAN>  
 A:Cross-references: EMBL:X61207; NID:938669; PIDD:CAA43515.1; PID:938670  
 C:Gene: hisB  
 C:Species: Azospirillum  
 C:Accession: J050045  
 C:Date: 07-Jun-1990 #sequence revision 07-Jun-1990 #text-change 31-Mar-2000  
 F:36-203/Domain: imidazoleglycerol-phosphate dehydratase homology <IPD>

Query Match 90.0%; Score 27; DB 2; Length 207;  
 Best Local Similarity 83.3%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSI/GCS 6  
 |:|||||  
 DB 201 GSI/GCS 206

Search completed: February 4, 2002, 08:00:55  
 Job time: 39 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1995 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:02:21 (Search time 35.81 seconds  
(without alignments))

24,508 Million cell updates/sec

Title: US09-642-660-11

Perfect score: 30

Sequence: 1 GSIQS 6

Scoring table: BIOSUM62

Gapop 10.0, Gapext 0.5

Searched: 473505 segs, 14672329 residues

Total number of hits satisfying chosen parameters: 473505

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database:

- 1: SP:archaea:\*
- 2: SP:bacteria:\*
- 3: SP:fungi:\*
- 4: SP:invertebrate:\*
- 5: SP:invertebrate:\*
- 6: SP:mammal:\*
- 7: SP:mammal:\*
- 8: SP:organelle:\*
- 9: SP:organelle:\*
- 10: SP:plant:\*
- 11: SP:plant:\*
- 12: SP:virus:\*
- 13: SP:virus:\*
- 14: SP:unclassified:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	30	100.0	107	1	OSY426 atropyrum p
2	30	100.0	107	1	OSY426 atropyrum p
3	30	100.0	107	1	OSY426 atropyrum p
4	30	100.0	209	12	OSY5018 human immun
5	30	100.0	277	5	OS61297 haemaphysal
6	30	100.0	281	2	OS4259 streptomyces
7	30	100.0	286	5	OSY460 streptomyces
8	30	100.0	371	2	OSY464 streptomyces
9	30	100.0	371	2	OSY464 streptomyces
10	30	100.0	750	10	OSY464 streptomyces
11	30	100.0	753	5	OSY464 streptomyces
12	30	100.0	773	10	OSY464 streptomyces
13	30	100.0	802	2	OSY464 streptomyces
14	30	100.0	888	2	OSY4125 actinobacill
15	30	100.0	898	2	OSY4125 actinobacill
16	30	100.0	931	2	OSY4127 actinobacill
17	30	100.0	931	2	OSY4127 actinobacill
18	30	100.0	935	3	OSY460 streptomyces
19	30	100.0	936	2	OSY460 streptomyces
20	30	100.0	936	2	OSY460 streptomyces
21	30	100.0	993	2	OSY426 atropyrum p
22	30	100.0	1013	3	OSY4127 actinobacill
23	30	100.0	1081	5	OSY464 streptomyces
24	30	100.0	1150	5	OSY464 streptomyces
25	30	100.0	1156	11	OSY464 streptomyces
26	30	100.0	1536	11	OSY464 streptomyces
27	30	100.0	1536	11	OSY464 streptomyces
28	30	100.0	2524	11	OSY464 streptomyces
29	30	100.0	2524	11	OSY464 streptomyces
30	30	100.0	2524	11	OSY464 streptomyces
31	30	100.0	2524	11	OSY464 streptomyces
32	30	100.0	2524	11	OSY464 streptomyces
33	30	100.0	2524	11	OSY464 streptomyces
34	30	100.0	2524	11	OSY464 streptomyces
35	30	100.0	2524	11	OSY464 streptomyces
36	30	100.0	2524	11	OSY464 streptomyces
37	30	100.0	2524	11	OSY464 streptomyces
38	30	100.0	2524	11	OSY464 streptomyces
39	30	100.0	2524	11	OSY464 streptomyces
40	30	100.0	2524	11	OSY464 streptomyces
41	30	100.0	2524	11	OSY464 streptomyces
42	30	100.0	2524	11	OSY464 streptomyces
43	30	100.0	2524	11	OSY464 streptomyces
44	30	100.0	2524	11	OSY464 streptomyces
45	30	100.0	2524	11	OSY464 streptomyces
46	30	100.0	2524	11	OSY464 streptomyces
47	30	100.0	2524	11	OSY464 streptomyces
48	30	100.0	2524	11	OSY464 streptomyces
49	30	100.0	2524	11	OSY464 streptomyces
50	30	100.0	2524	11	OSY464 streptomyces
51	30	100.0	2524	11	OSY464 streptomyces
52	30	100.0	2524	11	OSY464 streptomyces
53	30	100.0	2524	11	OSY464 streptomyces
54	30	100.0	2524	11	OSY464 streptomyces
55	30	100.0	2524	11	OSY464 streptomyces
56	30	100.0	2524	11	OSY464 streptomyces
57	30	100.0	2524	11	OSY464 streptomyces
58	30	100.0	2524	11	OSY464 streptomyces
59	30	100.0	2524	11	OSY464 streptomyces
60	30	100.0	2524	11	OSY464 streptomyces
61	30	100.0	2524	11	OSY464 streptomyces
62	30	100.0	2524	11	OSY464 streptomyces
63	30	100.0	2524	11	OSY464 streptomyces
64	30	100.0	2524	11	OSY464 streptomyces
65	30	100.0	2524	11	OSY464 streptomyces
66	30	100.0	2524	11	OSY464 streptomyces
67	30	100.0	2524	11	OSY464 streptomyces
68	30	100.0	2524	11	OSY464 streptomyces
69	30	100.0	2524	11	OSY464 streptomyces
70	30	100.0	2524	11	OSY464 streptomyces
71	30	100.0	2524	11	OSY464 streptomyces
72	30	100.0	2524	11	OSY464 streptomyces
73	30	100.0	2524	11	OSY464 streptomyces
74	30	100.0	2524	11	OSY464 streptomyces
75	30	100.0	2524	11	OSY464 streptomyces
76	30	100.0	2524	11	OSY464 streptomyces
77	30	100.0	2524	11	OSY464 streptomyces
78	30	100.0	2524	11	OSY464 streptomyces
79	30	100.0	2524	11	OSY464 streptomyces
80	30	100.0	2524	11	OSY464 streptomyces
81	30	100.0	2524	11	OSY464 streptomyces
82	30	100.0	2524	11	OSY464 streptomyces
83	30	100.0	2524	11	OSY464 streptomyces
84	30	100.0	2524	11	OSY464 streptomyces
85	30	100.0	2524	11	OSY464 streptomyces
86	30	100.0	2524	11	OSY464 streptomyces
87	30	100.0	2524	11	OSY464 streptomyces
88	30	100.0	2524	11	OSY464 streptomyces
89	30	100.0	2524	11	OSY464 streptomyces
90	30	100.0	2524	11	OSY464 streptomyces
91	30	100.0	2524	11	OSY464 streptomyces
92	30	100.0	2524	11	OSY464 streptomyces
93	30	100.0	2524	11	OSY464 streptomyces
94	30	100.0	2524	11	OSY464 streptomyces
95	30	100.0	2524	11	OSY464 streptomyces
96	30	100.0	2524	11	OSY464 streptomyces
97	30	100.0	2524	11	OSY464 streptomyces
98	30	100.0	2524	11	OSY464 streptomyces
99	30	100.0	2524	11	OSY464 streptomyces
100	30	100.0	2524	11	OSY464 streptomyces

RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amandlides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.

RT Courses: "  
RL J. Virol. 70:3589-3598(1996).  
DR EMBL/ U47563; AAB06904.1; -.



DR InterPro: IPR000777; GP120.  
 D1 EMBL: AF000161; GP120; 1.  
 K1 EMBL: AF000161; GP120; 1.  
 FT NON\_TER 203 203  
 SEQUENCE 203 AA; 22521 MW; B3FED7FC90B810EB CRC64;

Query Match 100.0%; Score 30; DB 12; Length 203;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 139 GSLOGS 144

## RESULT 4

ID 075018 PRELIMINARY: PRT; 209 AA.

DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ENVELOPE GLYCOPROTEIN (FRAGMENT).  
 CS Human immunodeficiency virus type 1.  
 OS Viruses; Retroviral viruses; Retroviridae; Lentivirus.  
 OX NCBI\_TaxId=1676;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TWIN A;  
 RA MEDLINE=96211491; PubMed=8648693;  
 RX PubMed=8648693; Author=Hill M, Scott M, Wood C.;  
 RT "Longitudinal studies of 'viral' sequences viral peptides  
 RT infection in perinatally infected twins with discordant disease  
 RT courses".  
 RL J. Virol. 70:3589-3596(1996).  
 DR EMBL: U47568; AAB06909.1;  
 DR PubMed=8648693; GP120.  
 DR EMBL: AF000161; GP120; 1.  
 FT Envelope protein. 1  
 K1 NON\_TER 1  
 FT NON\_TER 209 209  
 SEQUENCE 209 AA; 23290 MW; 00BDB584476FE65 CRC64;

Query Match 100.0%; Score 30; DB 12; Length 209;  
 Best Local Similarity 100.0%; Pred. No. 1.1e-00;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 144 GSLOGS 149

## RESULT 5

ID 061297 PRELIMINARY: PRT; 277 AA.

DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)  
 DE EXTRACELLULAR MATRIX-LIKE PROTEIN; COMPOSITE CDS.  
 OS Haemaphysalis longicornis.  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Acari;  
 OC Parasitiformes; Ixodida; Ixodidae; Haemaphysalis.  
 OX NCBI\_TaxId=44386;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=J133;  
 RA MEDLINE=96211491; PubMed=8648693;  
 RX PubMed=8648693; Author=Hill M, Scott M, Wood C.;  
 RT "Longitudinal studies of 'viral' sequences viral peptides  
 RT infection in perinatally infected twins with discordant disease  
 RT courses".  
 RL J. Virol. 70:3589-3596(1996).  
 DR EMBL: U47568; AAB06909.1;  
 DR PubMed=8648693; GP120.  
 DR EMBL: AF000161; GP120; 1.  
 FT Envelope protein. 1  
 K1 NON\_TER 1  
 FT NON\_TER 209 209  
 SEQUENCE 209 AA; 23290 MW; 00BDB584476FE65 CRC64;

DR EMBL: AB014612; BAA28620.1;  
 SQ SEQUENCE 277 AA; 27973 MW; 5B3297AE3A3BE09C CRC64;

Query Match 100.0%; Score 30; DB 5; Length 277;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 39 GSLOGS 44

## RESULT 6

ID 054259 PRELIMINARY: PRT; 281 AA.

AC 054259;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)  
 DE GLA 0 GENES STRB1, STRF, STRG, STRN, STRV.  
 CN STRF.  
 OS Streptomyces glaucocens.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomycetes.  
 OX NCBI\_TaxId=1907;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GLA 0;  
 RA MEDLINE=96211491; PubMed=8648693;  
 RX PubMed=8648693; Author=Hill M, Scott M, Wood C.;  
 RT "Longitudinal studies of 'viral' sequences viral peptides  
 RT infection in perinatally infected twins with discordant disease  
 RT courses".  
 RL J. Virol. 70:3589-3596(1996).  
 DR EMBL: U47568; AAB06909.1;  
 DR PubMed=8648693; GP120.  
 DR EMBL: AF000161; GP120; 1.  
 FT Envelope protein. 1  
 K1 NON\_TER 1  
 FT NON\_TER 209 209  
 SEQUENCE 209 AA; 23290 MW; 00BDB584476FE65 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 281;  
 Best Local Similarity 100.0%; Pred. No. 1.7e-02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 91 GSLOGS 96

## RESULT 7

ID 09VBD0 PRELIMINARY: PRT; 286 AA.

AC 09VBD0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)



05 Arabidopsis thaliana (Mouse-ear cress).  
 06 Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
 07 Spermatophyta: Magnoliophyta: eudicotyledons: core eudicots: Rosidae:  
 08 OC: Erythralia: Brassicales: Brassicaceae: Arabidopsis.  
 09 NCBI\_TaxID=3702.  
 10  
 11  
 12  
 13  
 14  
 15  
 16  
 17  
 18  
 19  
 20  
 21  
 22  
 23  
 24  
 25  
 26  
 27  
 28  
 29  
 30  
 31  
 32  
 33  
 34  
 35  
 36  
 37  
 38  
 39  
 40  
 41  
 42  
 43  
 44  
 45  
 46  
 47  
 48  
 49  
 50  
 51  
 52  
 53  
 54  
 55  
 56  
 57  
 58  
 59  
 60  
 61  
 62  
 63  
 64  
 65  
 66  
 67  
 68  
 69  
 70  
 71  
 72  
 73  
 74  
 75  
 76  
 77  
 78  
 79  
 80  
 81  
 82  
 83  
 84  
 85  
 86  
 87  
 88  
 89  
 90  
 91  
 92  
 93  
 94  
 95  
 96  
 97  
 98  
 99  
 100  
 101  
 102  
 103  
 104  
 105  
 106  
 107  
 108  
 109  
 110  
 111  
 112  
 113  
 114  
 115  
 116  
 117  
 118  
 119  
 120  
 121  
 122  
 123  
 124  
 125  
 126  
 127  
 128  
 129  
 130  
 131  
 132  
 133  
 134  
 135  
 136  
 137  
 138  
 139  
 140  
 141  
 142  
 143  
 144  
 145  
 146  
 147  
 148  
 149  
 150  
 151  
 152  
 153  
 154  
 155  
 156  
 157  
 158  
 159  
 160  
 161  
 162  
 163  
 164  
 165  
 166  
 167  
 168  
 169  
 170  
 171  
 172  
 173  
 174  
 175  
 176  
 177  
 178  
 179  
 180  
 181  
 182  
 183  
 184  
 185  
 186  
 187  
 188  
 189  
 190  
 191  
 192  
 193  
 194  
 195  
 196  
 197  
 198  
 199  
 200  
 201  
 202  
 203  
 204  
 205  
 206  
 207  
 208  
 209  
 210  
 211  
 212  
 213  
 214  
 215  
 216  
 217  
 218  
 219  
 220  
 221  
 222  
 223  
 224  
 225  
 226  
 227  
 228  
 229  
 230  
 231  
 232  
 233  
 234  
 235  
 236  
 237  
 238  
 239  
 240  
 241  
 242  
 243  
 244  
 245  
 246  
 247  
 248  
 249  
 250  
 251  
 252  
 253  
 254  
 255  
 256  
 257  
 258  
 259  
 260  
 261  
 262  
 263  
 264  
 265  
 266  
 267  
 268  
 269  
 270  
 271  
 272  
 273  
 274  
 275  
 276  
 277  
 278  
 279  
 280  
 281  
 282  
 283  
 284  
 285  
 286  
 287  
 288  
 289  
 290  
 291  
 292  
 293  
 294  
 295  
 296  
 297  
 298  
 299  
 300  
 301  
 302  
 303  
 304  
 305  
 306  
 307  
 308  
 309  
 310  
 311  
 312  
 313  
 314  
 315  
 316  
 317  
 318  
 319  
 320  
 321  
 322  
 323  
 324  
 325  
 326  
 327  
 328  
 329  
 330  
 331  
 332  
 333  
 334  
 335  
 336  
 337  
 338  
 339  
 340  
 341  
 342  
 343  
 344  
 345  
 346  
 347  
 348  
 349  
 350  
 351  
 352  
 353  
 354  
 355  
 356  
 357  
 358  
 359  
 360  
 361  
 362  
 363  
 364  
 365  
 366  
 367  
 368  
 369  
 370  
 371  
 372  
 373  
 374  
 375  
 376  
 377  
 378  
 379  
 380  
 381  
 382  
 383  
 384  
 385  
 386  
 387  
 388  
 389  
 390  
 391  
 392  
 393  
 394  
 395  
 396  
 397  
 398  
 399  
 400  
 401  
 402  
 403  
 404  
 405  
 406  
 407  
 408  
 409  
 410  
 411  
 412  
 413  
 414  
 415  
 416  
 417  
 418  
 419  
 420  
 421  
 422  
 423  
 424  
 425  
 426  
 427  
 428  
 429  
 430  
 431  
 432  
 433  
 434  
 435  
 436  
 437  
 438  
 439  
 440  
 441  
 442  
 443  
 444  
 445  
 446  
 447  
 448  
 449  
 450  
 451  
 452  
 453  
 454  
 455  
 456  
 457  
 458  
 459  
 460  
 461  
 462  
 463  
 464  
 465  
 466  
 467  
 468  
 469  
 470  
 471  
 472  
 473  
 474  
 475  
 476  
 477  
 478  
 479  
 480  
 481  
 482  
 483  
 484  
 485  
 486  
 487  
 488  
 489  
 490  
 491  
 492  
 493  
 494  
 495  
 496  
 497  
 498  
 499  
 500  
 501  
 502  
 503  
 504  
 505  
 506  
 507  
 508  
 509  
 510  
 511  
 512  
 513  
 514  
 515  
 516  
 517  
 518  
 519  
 520  
 521  
 522  
 523  
 524  
 525  
 526  
 527  
 528  
 529  
 530  
 531  
 532  
 533  
 534  
 535  
 536  
 537  
 538  
 539  
 540  
 541  
 542  
 543  
 544  
 545  
 546  
 547  
 548  
 549  
 550  
 551  
 552  
 553  
 554  
 555  
 556  
 557  
 558  
 559  
 560  
 561  
 562  
 563  
 564  
 565  
 566  
 567  
 568  
 569  
 570  
 571  
 572  
 573  
 574  
 575  
 576  
 577  
 578  
 579  
 580  
 581  
 582  
 583  
 584  
 585  
 586  
 587  
 588  
 589  
 590  
 591  
 592  
 593  
 594  
 595  
 596  
 597  
 598  
 599  
 600  
 601  
 602  
 603  
 604  
 605  
 606  
 607  
 608  
 609  
 610  
 611  
 612  
 613  
 614  
 615  
 616  
 617  
 618  
 619  
 620  
 621  
 622  
 623  
 624  
 625  
 626  
 627  
 628  
 629  
 630  
 631  
 632  
 633  
 634  
 635  
 636  
 637  
 638  
 639  
 640  
 641  
 642  
 643  
 644  
 645  
 646  
 647  
 648  
 649  
 650  
 651  
 652  
 653  
 654  
 655  
 656  
 657  
 658  
 659  
 660  
 661  
 662  
 663  
 664  
 665  
 666  
 667  
 668  
 669  
 670  
 671  
 672  
 673  
 674  
 675  
 676  
 677  
 678  
 679  
 680  
 681  
 682  
 683  
 684  
 685  
 686  
 687  
 688  
 689  
 690  
 691  
 692  
 693  
 694  
 695  
 696  
 697  
 698  
 699  
 700  
 701  
 702  
 703  
 704  
 705  
 706  
 707  
 708  
 709  
 710  
 711  
 712  
 713  
 714  
 715  
 716  
 717  
 718  
 719  
 720  
 721  
 722  
 723  
 724  
 725  
 726  
 727  
 728  
 729  
 730  
 731  
 732  
 733  
 734  
 735  
 736  
 737  
 738  
 739  
 740  
 741  
 742  
 743  
 744  
 745  
 746  
 747  
 748  
 749  
 750  
 751  
 752  
 753  
 754  
 755  
 756  
 757  
 758  
 759  
 760  
 761  
 762  
 763  
 764  
 765  
 766  
 767  
 768  
 769  
 770  
 771  
 772  
 773  
 774  
 775  
 776  
 777  
 778  
 779  
 780  
 781  
 782  
 783  
 784  
 785  
 786  
 787  
 788  
 789  
 790  
 791  
 792  
 793  
 794  
 795  
 796  
 797  
 798  
 799  
 800  
 801  
 802  
 803  
 804  
 805  
 806  
 807  
 808  
 809  
 810  
 811  
 812  
 813  
 814  
 815  
 816  
 817  
 818  
 819  
 820  
 821  
 822  
 823  
 824  
 825  
 826  
 827  
 828  
 829  
 830  
 831  
 832  
 833  
 834  
 835  
 836  
 837  
 838  
 839  
 840  
 841  
 842  
 843  
 844  
 845  
 846  
 847  
 848  
 849  
 850  
 851  
 852  
 853  
 854  
 855  
 856  
 857  
 858  
 859  
 860  
 861  
 862  
 863  
 864  
 865  
 866  
 867  
 868  
 869  
 870  
 871  
 872  
 873  
 874  
 875  
 876  
 877  
 878  
 879  
 880  
 881  
 882  
 883  
 884  
 885  
 886  
 887  
 888  
 889  
 890  
 891  
 892  
 893  
 894  
 895  
 896  
 897  
 898  
 899  
 900  
 901  
 902  
 903  
 904  
 905  
 906  
 907  
 908  
 909  
 910  
 911  
 912  
 913  
 914  
 915  
 916  
 917  
 918  
 919  
 920  
 921  
 922  
 923  
 924  
 925  
 926  
 927  
 928  
 929  
 930  
 931  
 932  
 933  
 934  
 935  
 936  
 937  
 938  
 939  
 940  
 941  
 942  
 943  
 944  
 945  
 946  
 947  
 948  
 949  
 950  
 951  
 952  
 953  
 954  
 955  
 956  
 957  
 958  
 959  
 960  
 961  
 962  
 963  
 964  
 965  
 966  
 967  
 968  
 969  
 970  
 971  
 972  
 973  
 974  
 975  
 976  
 977  
 978  
 979  
 980  
 981  
 982  
 983  
 984  
 985  
 986  
 987  
 988  
 989  
 990  
 991  
 992  
 993  
 994  
 995  
 996  
 997  
 998  
 999  
 1000

Query Match 100.0% Score 30; DB 10; Length 750;  
 Best Local Similarity 100.0%; Pred. No. 5.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Db 733 GSLOGS 738

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13)  
 DE 01-MAY-2000 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL 78.7 KDA PROTEIN.  
 OS Leishmania major.  
 NC Eukaryota: Euklenozoa: Kinetoplastida: Trypanosomatidae: Leishmania.  
 NC NCBI\_TaxID=5684.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC Straub-Friedlin.  
 RP SEQUENCE FROM N.A.  
 RA Straub-Friedlin.  
 RC MEDLINE-98146435: PubMed-9477341.  
 RA Iyena A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
 Smith D.F.;  
 RT "A physical map of the Leishmania major Friedlin genome.";  
 RC Genome Res. 13:15-18(1999).  
 RW PubMed-9477341: PubMed-9551513: -.  
 DE Hypochemical protein.  
 SO SEQUENCE 763 AA; 78737 MW; 815N15ECAA1265361 CRC64.

090193  
 ID 090193 PRELIMINARY; PRT; 763 AA.  
 RC 090193-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (T

OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=6036198; PubMed=7582000;  
 RA Gonzalez G.C., Yu R.H., Roelack P.R. Jr., Schryvers A.B.;  
 RT "Sequence, genetic analysis, and expression of Actinobacillus  
 pleuropneumoniae transferrin receptor genes";  
 RL Microbiology 141:2405-2416(1995).  
 DR InterPro: IPR005931; TOMB\_BOXC.  
 DR Pfam: PF00593; TOMB\_BOXC.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 SQ SEQUENCE 898 AA; 102187 MW; 484968810A3752A CRC64;

Query Match 100.0%; Score 30; DB 2; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOGS 6  
 DB 151 GSLOGS 156

RESULT 15  
 OX04158 PRELIMINARY; PRT: 898 AA.  
 AC 044158  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE TRANSFERRIN RECEPTOR.  
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=6036198;  
 RA Mille M.;  
 RL Tieraerztliche Hochschule Hannover, Germany.  
 DR EMBL: Z54191; CA90896.1;  
 DR InterPro: IPR005931; TOMB\_BOXC.  
 DR Pfam: PF00593; TOMB\_BOXC.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 SQ SEQUENCE 898 AA; 102217 MW; A06275B0C0B76E52 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 898;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOGS 6  
 DB 151 GSLOGS 156

RESULT 16  
 OX04127 PRELIMINARY; PRT: 931 AA.  
 AC 044127  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE TRANSFERRIN BINDING PROTEIN 1.  
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
 OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=6036198; PubMed=7582000;  
 RA Gonzalez G.C., Yu R.H., Roelack P.R. Jr., Schryvers A.B.;  
 RT "Sequence, genetic analysis, and expression of Actinobacillus  
 pleuropneumoniae transferrin receptor genes";  
 RL Microbiology 141:2405-2416(1995).  
 DR InterPro: IPR005931; TOMB\_BOXC.  
 DR Pfam: PF00593; TOMB\_BOXC.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 SQ SEQUENCE 931 AA; 106326 MW; D77306E22DA54F9 CRC64;

OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=6036198; PubMed=7582000;  
 RA Gonzalez G.C., Yu R.H., Roelack P.R. Jr., Schryvers A.B.;  
 RT "Sequence, genetic analysis, and expression of Actinobacillus  
 pleuropneumoniae transferrin receptor genes";  
 RL Microbiology 141:2405-2416(1995).  
 DR InterPro: IPR005931; TOMB\_BOXC.  
 DR Pfam: PF00593; TOMB\_BOXC.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 SQ SEQUENCE 931 AA; 106369 MW; C688D077D03C54F9 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 931;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOGS 6  
 DB 151 GSLOGS 156

RESULT 17  
 OX044166 PRELIMINARY; PRT: 931 AA.  
 AC 044166  
 DT 01-NOV-1996 (TEMBLrel. 01, Created)  
 DT 01-NOV-1996 (TEMBLrel. 01, Last sequence update)  
 DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)  
 DE TBP1 GENE PREDICTION.  
 OS Actinobacillus pleuropneumoniae (Haemophilus pleuropneumoniae).  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Actinobacillus.  
 OX NCBI\_TaxID=715;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 27088;  
 RX MEDLINE=96207589; PubMed=8670116;  
 RA Daban M., Medrano A., Queiro E.;  
 RT "Cloning, sequencing and expression of the transferrin-binding protein  
 1 gene from Actinobacillus pleuropneumoniae";  
 RL Biochem J. 315:257-264(1996).  
 DR EMBL: Z64918; CA900531; TOMB\_BOXC.  
 DR InterPro: IPR005931; TOMB\_BOXC.  
 DR Pfam: PF00593; TOMB\_BOXC.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 DR PROSITE: PS01156; TOMB\_DEPENDENT\_REC\_2; 1.  
 KW Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 931 TBP1 GENE.  
 FT SEQUENCE 931 AA; 106326 MW; D77306E22DA54F9 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 931;  
 Best Local Similarity 100.0%; Pred. No. 6.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GSLOGS 6  
 DB 151 GSLOGS 156

RESULT 18  
 OX09620 PRELIMINARY; PRT: 935 AA.  
 ID 09620  
 DT 01-OCT-2000 (TEMBLrel. 15, Created)  
 DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)

```

DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DR CONSERVED HYPOTHEICAL PROTEIN.
DN 3511.260
NC Nucleoside
OC Bacteriota: Fungi: Ascomycota: Pezizomycotina: Sordariomycetes;
OX Sordariiales; Sordariaceae; Neurospora.
RN [1]
RA NCBJ_TaxID=5141;
RP SEQUENCE FROM N.A.
RA Schulte U., Algin V., Hohenisel J., Brandt P., Fartmann B., Holland R.,
RA Mykura G., Meves H.W., Mennhapf G.;
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
RP [2]
RP SEQUENCE FROM N.A.
RA German Neurospora genome project;
RA Submitted (Apr-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AL353820; CAB8594.1; -
DR InterPro: IPR000822; znf_C2H2.
DR SMART: SM00355; z1-C2H2_2.
DR PROSITE: PS00026; ZINC_FINGER_C2H2_1; 2.
DR PROSITE: PS01057; ZINC_FINGER_C2H2_2; 2.
RM DNA-binding; Metal-binding; Zinc-finger.
SQ SEQUENCE 935 AA; 103980 MW; 35F06450AC2BD20 CRC64;

Query Match 100.0%; Score 30; DB 3; Length 935;
Best Local Similarity 100.0%; Pred No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLGGS 6
DB 700 GSLGGS 705

RESULT 19
Q9RZS3 PRELIMINARY; PRT; 936 AA.
ID 09RZS3
AC 09RZS3
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 19-2001 (TREMBLrel. 13, Last annotation update)
DE HYPOTHEICAL 97.7 KDa PROTEIN.
CN DR00041
OS Deinococcus radiodurans.
OC Plasmid MP1.
OC Bacteria: Thermus/Deinococcus group; Deinococcales; Deinococcus.
OX NCBJ_TaxID=1299;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN=RI.
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Melillo O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,
RA Modfart K.S., Qin H., Jiang L., Pauphille W., Crosby M., Shen M.,
RA Khatkhatya K.S., Alayand L., Daly M.J., Milton K.W., Fleischmann R.D.,
RA Fraser C.M., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.;
RA *Genome sequence of the radioresistant bacterium Deinococcus
RA radiodurans RI.*;
RL Science 286:1571-1577(1999).
DR EMBL: AE001826; AAF12631.1; -
DR TIGR: DR00041;
DR PROSITE: PS00173; PFRB
DR PROSITE: PS00583; PFRB_XINASES_1; UNKNOWN_1
RM Hypothetical protein; Plasmid; Complete proteome.
SQ SEQUENCE 936 AA; 97723 MW; 4ABD43D7F195C0 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 936;
Best Local Similarity 100.0%; Pred No. 6.9e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GSLGGS 6
DB 800 GSLGGS 805

RESULT 20
Q9KIV0 PRELIMINARY; PRT; 993 AA.
ID 09KIV0
AC 09KIV0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HEMOGLOBIN BINDING PROTEIN C.
CN HSCB
OS Haemophilus influenzae.
OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBJ_TaxID=727;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN=N182.
RA MEDLINE=20316037; PubMed=10659226;
RC Cope L.D., Hkrel Z., Hansen E.J.;
RT "Detection of Phase Variation in Expression of Proteins Involved in
RT Haemophilus and Hemoglobin-Haptoglobin Binding by Nontypeable
RT Haemophilus influenzae.";
RL Infect. Immun. 68:4097-4101(2000).
DR EMBL: AF2180; A48697.1; -
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00593; TonB_box; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.
SQ SEQUENCE 993 AA; 113616 MW; A51BF3B2C5441612 CRC64;

Query Match 100.0%; Score 30; DB 2; Length 993;
Best Local Similarity 100.0%; Pred No. 7.4e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLGGS 6
DB 174 GSLGGS 179

RESULT 21
O87296 PRELIMINARY; PRT; 999 AA.
ID 087296
AC 087296;
DT 01-NOV-1998 (TREMBLrel. 08, Created)
DT 01-NOV-1998 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE HEMOGLOBIN BINDING PROTEIN.
CN HGFB.
OS Haemophilus influenzae.
OC Bacteria: Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Haemophilus.
OX NCBJ_TaxID=727;
RN [1]
RA SEQUENCE FROM N.A.
RP STRAIN=H1689 TYPE B.
RA MEDLINE=98427137; PubMed=9746572;
RA Ren Z., Jin H., Morton D.J., Stull T.L.;
RT "HspA, a gene encoding a second Haemophilus influenzae hemoglobin- and
RT hemoglobin-haptoglobin-binding protein.";
RL Infect. Immun. 66:4339-4341(1998).
DR EMBL: AF022910; 66:4339-4341(1998).
DR InterPro: IPR000531; TonB_box;
DR InterPro: IPR001993; Mitoch_carrier.
DR Pfam: PF00593; TonB_box; 1.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_1.
DR PROSITE: PS01156; TONB_DEPENDENT_REC_2; 1.

```



OC	Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:
OC	Phylogeny: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:
OC	Ephyridae: Drosophilidae: Drosophila.
OC	NCBI_TaxID=7227.
CP	SEQUENCE FROM N.A.
RP	STRAIN=BERKELEY.
RX	MEDLINE=20196006; PUBMED=10731132;
RA	Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA	Angelidis P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA	George K.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA	Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA	Man K.H., McQuinn K., Dvorkin I., Doyle J.K., Batzli G.L.,
RA	Appel J.F., Asyran A.A., An H.-J., Andrews-Platch C., Matsuyama G.,
RA	Bailey R.M., Bayraktar A., An H.-J., Bayraktar-Platch L., Beasley E.M.,
RA	Beeoon K.Y., Benos P.V., Benham B.P., Bhandari D., Bolshakov S.,
RA	Berova D., Botchan M.R., Butler H., Brockstein P., Brothier P.,
RA	Burtis K.C., Busam D.A., Butler H., Cardout E., Center A., Chandra I.,
RA	DePalma D., Delaney S., Dhallin C., Davenport L.B., Davies P.,
RA	Dodson K., Dong L.E., Downes M., Dugan-Rubio S., Dunoyevich S.,
RA	Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA	Foster C., Gabrielian A.E., Garg N.Z., Gelbart W.M., Glasser K.,
RA	Giodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA	Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA	Hosdin D., Houston K.A., Howard T.J., Mei M.-H., Ibeagwa C.,
RA	Kimmel P.M., Klotzel K., Kropfen K.H., Ke Z., Krennson J.A., Ketchum K.A.,
RA	Kimball J.R., Lasko P.M., LeVay S.A., Li J., Litz J., Lupu D., Lai Z.,
RA	Liu X., Maitel B., Melnick N.V., McBratney C., Melrod M.P., McPherson D.A.,
RA	Merrillow G., Milhina N.V., Mohrty C., Morris C., Morrison J.,
RA	Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA	Nelson D.R., Nelson K.S., Nixon K., Nisikawa D.R., Paclet J.M.,
RA	Paezoldo M., Pitman K.S., Pan S., Pollard J., Puti V., Reese M.G.,
RA	Shen B.P., Shengdong K., Saunders R.D.C., Scheeler F., Shen H.,
RA	Spier E.C., Spreading A.C., Stimpelman M., Styrsky M.F., Smith T.,
RA	Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA	Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA	Williams Z., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA	Ye Y., Yeh R.-F., Yavel J.S., Zhou M., Zhang G., Zhao Q., Zhang L.,
RA	Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA	Yih-Hsien Wu, Byers E.M., Rubin G.M., Venter J.C.,
RT	7th International Human Genome Conference, 1996.
RT	Science 287:2185-2193(1999).
DR	EMBL: A003678; A055239.2.
DR	PubMed: 7990037566; C0111729.
SO	SEQUENCE 1340 AA; 147657 MW; 8E2AC536CA0CAED CRC64;

RP SEQUENCE FROM N. A.  
RA STRAHL-FRIEDLID;  
RA STRAHL-FRIEDLID; Hixson G., Lemley C., Ricel E.,  
RA Sisk E., Shukla S., Swartzliff S., Kistler T., Magness C., Bastien P.,  
RA F.G., Ivens A., Stuart R.;  
RA The amino acid sequence of chresosome 1 has only two polyclonistic  
units of protein coding information.  
RN Submitted (May-1998) to the EMBL/genBank/DDB databases.  
RL [2]  
RP SEQUENCE FROM N. A.  
RP STRAHL-FRIEDLID;  
RP STRAHL-FRIEDLID;  
PT The nucleotide sequence of Leishmania major Friedlin chromosome 1.\*;  
PT Submitted (May-1998) to the EMBL/genBank/DDB databases.  
CC -1- SIMILARITY: TO THE SER/THE FAMILY OF PROTEIN KINASES.  
EMBL: AE001274; AAC24682.1;  
InterPro: IPR000719; Ser\_Pkinase.  
DR Pfam PF00069;G02290; Ser\_Pkinase.  
DR Pfam PF00069;G02290; Ser\_Pkinase.  
DR POSITIVE: PSS0101; PROTEIN\_KINASE\_DOM; 1.  
DR POSITIVE: PSS0108; PROTEIN\_KINASE\_ST\_1.  
DR ATP-binding; serine/threonine-protein kinase; transferase.  
KW SEQUENCE 1557 AA; 167982 MW; RC69561AB8D086A CRC64;

```

Query Match Similarity      100.0%; Score 30; DB 5; Length 1557;
Best Local Similarity      100.0%; Pred. No. 1.2e+03;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0
OY      1 GUSTACS 6
      111111
Db      158 GUSTACS 163

RESULT      26
ID      062766      PRELIMINARY;      PTR; 1596 AA.
AC      062766;
DT      01-NOV-1996 (TEMBUREL; 01, Created)
DI      01-NOV-1996 (TEMBUREL; 01, Last sequence update)
DE      01-NOV-2000 (TEMBUREL; 17, Last annotation update)
DE      SERRKS.
GN      332.
OS      Rattus norvegicus (Rat).
OC      Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;
OC      Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX      NCBI_TaxID=10116;
RN      1)
RP      SEQUENCE FROM N.A.
RC      STRAIN=FISHER; TISSUE=EMBRIO;
RX      MEDLINE=95325937; PUBMED=739536;
RA      Gelman, T.; Nelson P.C.; Frankfort B.; Tomblar E.; Johnson R.;
RT      "Isolation and characterization of a novel mitogenic regulatory gene,
RT      322, which is transcriptionally suppressed in cells transformed by src
RT      and ras."
RL      Mol. Cell. Biol. 15:2754-2762(1995).
DR      EMBL; U23146; AAA79517.1; -.
DR      Interpro: IPR001573; PfamA:anrch.
DR      SEQUENCE: 1596 AA; 172498 MW; C678BD2161D9C7E2 CRC64;

```

```

Query Match      100.0%  Score 30:  DB 11:  Length 1566;
Best Local Similarity 100.0%;  Pred. No. 1,36+03;
Matches      6;  Conservative      0;  Mismatches      0;  Indels      0;  Caps      0;

QY      1  GSIQSS 6
          011111
Db      1347  GSIQSS 1352

RESULT      27
092117

```





MEMLINE-21016719; PubMed-11130712;  
 RA MEDLER A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White R., Altschul S., Chao A., Arzoo R., Bowman C.L., Brooks S.Y.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Coss G.W.,  
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hulzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.-J., Koo H.L., Kremesko S., Kutz D.B., Kwan A., Lam B.,  
 RA Lin S., Liu S., Lee A., Lee J.M., Lenz G.A., Li J.H., Li Y.-P.,  
 RA Maltsev S., Miranda M., Nguyen M., S. Halli R., Marshall A.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rooley D.,  
 RA Sano H., Salzer S.L., Schwartz J.R., Shlan P., Southwick A.,  
 RA Sun H., Tallon L.J., Tambaung G., Toriumi M.J., Tom C.D.,  
 RA Utterback T., Van Aken S., Vaynsberg M., Vysotskaya V.S., Walker M.,  
 RA \*Sequencing and analysis of chromosome 1 of the plant Arabidopsis  
 PT thaliana. 408:816-820(2000).  
 RL Nature 408:816-820(2000).  
 DR EMBL: AC079287; AAC50846.1; -  
 SO SEQUENCE 211 AA; 23856 MW; 728076C9AF285AB CRC64;

Query Match 93.3%; Score 28; DB 10; Length 211;  
 Best Local Similarity 83.3%; Pred. No. 3.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 150 GSLOGS 155

RESULT 31  
 0912B5 PRELIMINARY: PRT: 252 AA.  
 ID 0912B5  
 AC 0912B5  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last annotation update)  
 DE HYPOHETICAL 26.4 KDa PROTEIN.  
 GN F8P6.160.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 CC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
 CC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 RN [1] TaxID=9102;  
 RP SEQUENCE FROM N.A.  
 RA Bevan M., Peters S.A., van Staveren M., Dirks W., Stiekema W.,  
 RA Bancroft I., Mewes H.W., Rudd S., Lemke K., Meyer K.F.X.,  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project.  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL62873; CAB85515.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 252 AA; 28426 MW; 83C716AFB88866 CRC64;

Query Match 93.3%; Score 28; DB 10; Length 252;  
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSLOGS 6  
 Db 154 GSLOGS 159  
 RESULT 32  
 09FLK2 PRELIMINARY: PRT: 254 AA.

AC 09FLK2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE HYPOHETICAL PROTEIN TAO226.  
 OS Thermoplasma acidophilum.  
 CC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
 CC Thermoplasma  
 RN [1] TaxID=2303;  
 RP SEQUENCE FROM N.A.  
 RA MEDLER A., Becker J.R., Palm C.J., Federspiel N.A., Kaul S.,  
 RA White R., Altschul S., Chao A., Arzoo R., Bowman C.L., Brooks S.Y.,  
 RA Chung M.K., Conn L., Conway A.B., Conway A.R., Coss G.W.,  
 RA Dunn P., Egu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,  
 RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Hulzer L.,  
 RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,  
 RA Kim C.-J., Koo H.L., Kremesko S., Kutz D.B., Kwan A., Lam B.,  
 RA Lin S., Liu S., Lee A., Lee J.M., Lenz G.A., Li J.H., Li Y.-P.,  
 RA Maltsev S., Miranda M., Nguyen M., S. Halli R., Marshall A.,  
 RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rooley D.,  
 RA Sano H., Salzer S.L., Schwartz J.R., Shlan P., Southwick A.,  
 RA Sun H., Tallon L.J., Tambaung G., Toriumi M.J., Tom C.D.,  
 RA Utterback T., Van Aken S., Vaynsberg M., Vysotskaya V.S., Walker M.,  
 RA \*The genome sequence of the thermophilic scavenger Thermoplasma  
 AC acidophilum. 407:508-513(2000).  
 RL Nature 407:508-513(2000).  
 DR EMBL: AL445063; CAC1371.1; -  
 SO SEQUENCE 254 AA; 28522 MW; 241FC1271812031 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 254;  
 Best Local Similarity 83.3%; Pred. No. 3.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GSLOGS 6  
 Db 73 GSLOGS 78

RESULT 33  
 053815 PRELIMINARY: PRT: 267 AA.  
 ID 053815  
 AC 053815  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last annotation update)  
 DE ISP 5564 GENES STRB AND STRF (F8ACQ0MNT).  
 GN STRF.  
 CC Bacteria; Firmicutes; Actinobacteriia; Actinobacteridae;  
 CC Bacillales; Bacillaceae; Streptococcaceae; Streptococcus.  
 RN [1] TaxID=13897;  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-ISP 5564;  
 RA Meyer E.E., Pleschkeberg M.;  
 RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: X78972; CAA55568.1; -  
 FT NON-TER 267  
 SO SEQUENCE 267 AA; 29777 MW; D0D25243D6AC3B0 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 267;  
 Best Local Similarity 83.3%; Pred. No. 4.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSLOGS 6  
 Db 91 GSLOGS 96  
 RESULT 34  
 09FL20 PRELIMINARY: PRT: 280 AA.  
 ID 09FL20  
 AC 09FL20  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
 DE SIMILARITY TO MUTATOR-LIKE TRANSPOSASE.

OS Aradidopsals thaliana (Mouse-ear cress).  
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
 OC Spermatophyta: Magnoliophyta: eudicotyledons: core eudicotyledons: Rosidae:  
 OC Ericaceae: Iliaceae: Brassicales: Brassicaceae: Arabidopsals.  
 RN [1]\_TaxID=3702;  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=COLUMBIA;  
 RC MEDLINE=98344145; PubMed=9679202;  
 RA Tanaka T., Kotani H., Nakamura Y., Sato S., Asamizu E., Miyajima N.,  
 RT The genome sequence of Arabidopsis thaliana chromosome 5. V. Sequence  
 RT features of the regions of 1,381,566 bp covered by twenty one  
 RT physically assigned P1 and TAC clones."  
 RL DNA Res. 5:131-145(1998).  
 DR EMBL: AB010697; BAB10892.1;  
 SO SEQUENCE 280 AA: 31489 MW: 852051651DD1372 CRC64;

Query Match 93.3%; Score 28; DB 10; Length 280;  
 Best Local Similarity 83.3%; Pred. No. 4; e=02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSLOGS 6  
 Db 154 GSLOGS 159

RESULT 35  
 ID 09VBZ2 PRELIMINARY; PRT; 288 AA.  
 AC GSVEZ2 2000 (TREMBLrel. 13, Created)  
 DT 01-NOV-2000 (TREMBLrel. 13, Last annotation update)  
 DT 01-MAY-2000 (EMBLrel. 13, Last annotation update)  
 DE CG5468 PROTEIN.  
 GN CG5468  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:  
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:  
 OC Cyclorhina: Drosophilidae: Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Man K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Brill J.F., Adayelani A., An H.-J., Andrews-Pfannkoch L., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
 RA Borovica D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
 RA DeCarolis B., Delcher A., Deng Z., Davis A.C., Davenport L.B., Davies P.,  
 RA Cherry J.M., Chewey S., Dahlke C., Davenport L.B., Davies P., Chandra I.,  
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunbar B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostettler D., Houston K.A., Howland T.J., Wei M.-H., Iyegyan C.,  
 RA Jaisl M., Kalush P., Karpis G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X., Maitel B., Maitel C., Maitel A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., Maitel C., Maitel A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Mouton S.M., Moy M., Murphy K., Murphy L., Murry D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pachter J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,  
 RA Spitzer E., Spitzing A.C., Stapleton M., Strong R., Sun E.,

RA Relbett K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klimos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spitzing A.C., Stapleton M., Strong R., Sun E.,  
 RA Stryker E., Stryker E., Stryker E., Stryker E., Stryker E.,  
 RA Williams S.M., Woodage T., Morley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yen R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.W., Venter J.C.;  
 RT The genome sequence of Drosophila melanogaster."  
 RL Science 287:1228-1235(2000).  
 DR EMBL: AB010697; BAB10892.1;  
 SO SEQUENCE 288 AA: 28669 MW: AC2BBD3BDF18056 CRC64;

Query Match 93.3%; Score 28; DB 5; Length 288;  
 Best Local Similarity 83.3%; Pred. No. 4; e=02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSLOGS 6  
 Db 69 GSLOGS 74

RESULT 36  
 ID 09VBZ2 PRELIMINARY; PRT; 392 AA.  
 AC 077475;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DT 01-NOV-2001 (TREMBLrel. 17, Last annotation update)  
 DE CG5468 PROTEIN.  
 GN CG5468  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta:  
 OC Pterygota: Neoptera: Endopterygota: Diptera: Brachycera: Muscomorpha:  
 OC Cyclorhina: Drosophilidae: Drosophila.  
 NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RC MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA Sutton G.G., Morten J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Man K.H., Doyle G., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Brill J.F., Adayelani A., An H.-J., Andrews-Pfannkoch L., Baldwin D.,  
 RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
 RA Borovica D., Botchan M.R., Bouck J., Brockstein P., Brothier P.,  
 RA DeCarolis B., Delcher A., Deng Z., Davis A.C., Davenport L.B., Davies P.,  
 RA Cherry J.M., Chewey S., Dahlke C., Davenport L.B., Davies P., Chandra I.,  
 RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunbar B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Gary N., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Hostettler D., Houston K.A., Howland T.J., Wei M.-H., Iyegyan C.,  
 RA Jaisl M., Kalush P., Karpis G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X., Maitel B., Maitel C., Maitel A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Maitel B., Maitel C., Maitel A.C., Li J., Li Z., Liang Y., Lin X.,  
 RA Mouton S.M., Moy M., Murphy K., Murphy L., Murry D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pachter J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Put V., Reese M.G.,  
 RA Spitzer E., Spitzing A.C., Stapleton M., Strong R., Sun E.,





---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OK Protein - protein search, using sw model

Run on: February 4, 2002, 08:01:36 ; Search time 19.55 Seconds  
(without alignments)  
6,768 Million cell updates/sec

Title: US-09-642-660-11  
Perfect score: 130  
Sequence: 1 GSIACS 6

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 212252 seqs, 2250392 residues

Total number of hits satisfying chosen parameters: 212252

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 100 summaries

Database :  
1: /cgn2\_6/pdata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/pdata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/pdata/2/1aa/5C.COMB.pep.\*  
4: /cgn2\_6/pdata/2/1aa/5D.COMB.pep.\*  
5: /cgn2\_6/pdata/2/1aa/5E.COMB.pep.\*  
6: /cgn2\_6/pdata/2/1aa/5F.COMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	100.0	6	3	US-08-628-112-11	Sequence 11, Appl
2	100.0	6	3	US-08-628-112-11	Sequence 11, Appl
3	100.0	1346	2	US-08-613-125-11	Sequence 2, Appl
4	28	93.3	21	US-08-014-153D-45	Sequence 45, Appl
5	28	93.3	633	US-09-041-991A-8	Sequence 8, Appl
6	27	90.0	163	US-09-041-991A-10	Sequence 10, Appl
7	27	90.0	163	US-08-704-331-6	Sequence 6, Appl
8	27	90.0	246	US-08-704-331-2	Sequence 4, Appl
9	27	90.0	254	US-09-751-511-4	Sequence 4, Appl
10	27	90.0	318	US-08-633-148-4	Sequence 8, Appl
11	27	90.0	318	US-08-633-148-2	Sequence 2, Appl
12	27	90.0	326	US-08-633-543-2	Sequence 2, Appl
13	27	90.0	340	US-08-633-148-2	Sequence 2, Appl
14	27	90.0	349	US-08-633-172-6	Sequence 6, Appl
15	27	90.0	349	US-08-633-172-6	Sequence 6, Appl
16	27	90.0	349	US-08-633-172-6	Sequence 6, Appl
17	27	90.0	349	US-08-633-172-6	Sequence 6, Appl
18	27	90.0	513	US-08-646-514-2	Sequence 2, Appl
19	27	90.0	513	US-08-646-514-2	Sequence 2, Appl
20	27	90.0	624	US-08-646-514-2	Sequence 2, Appl
21	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
22	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
23	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
24	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
25	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
26	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
27	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
28	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
29	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
30	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
31	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
32	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
33	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
34	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
35	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
36	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
37	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
38	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
39	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
40	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
41	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
42	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
43	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
44	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
45	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
46	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
47	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
48	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
49	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
50	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
51	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
52	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
53	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
54	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
55	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
56	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
57	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
58	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
59	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
60	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
61	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
62	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
63	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
64	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
65	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
66	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
67	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
68	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
69	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
70	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
71	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
72	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
73	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
74	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
75	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
76	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
77	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
78	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
79	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
80	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
81	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
82	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
83	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
84	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
85	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
86	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
87	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
88	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
89	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
90	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
91	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
92	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
93	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
94	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
95	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
96	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
97	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
98	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
99	27	90.0	773	US-08-646-514-2	Sequence 2, Appl
100	27	90.0	773	US-08-646-514-2	Sequence 2, Appl

28	27	90.0	1000	2	US-08-867-541-12	Sequence 12, Appl
29	27	90.0	1000	2	US-08-867-541-12	Sequence 12, Appl
30	27	90.0	1000	4	US-09-074-658-12	Sequence 12, Appl
31	27	90.0	1000	4	US-09-074-658-12	Sequence 12, Appl
32	27	50.0	1342	2	US-08-484-438-9	Sequence 9, Appl
33	27	50.0	2432	4	US-09-074-658-15	Sequence 15, Appl
34	27	50.0	2432	4	US-09-074-658-15	Sequence 15, Appl
35	26	86.7	15	1	US-07-609-716-67	Sequence 16, Appl
36	26	86.7	15	1	US-07-609-716-67	Sequence 16, Appl
37	26	86.7	15	1	US-08-477-0598-37	Sequence 37, Appl
38	26	86.7	15	3	US-08-482-0889-37	Sequence 37, Appl
39	26	86.7	15	4	US-08-475-111A-46	Sequence 15, Appl
40	26	86.7	15	4	US-08-475-111A-46	Sequence 15, Appl
41	26	86.7	15	4	US-08-478-029A-16	Sequence 67, Appl
42	26	86.7	15	4	US-08-478-029A-16	Sequence 67, Appl
43	26	86.7	48	1	US-08-063-472B-46	Sequence 46, Appl
44	26	86.7	59	1	US-07-609-716-83	Sequence 83, Appl
45	26	86.7	59	4	US-08-475-111A-83	Sequence 83, Appl
46	26	86.7	59	4	US-08-475-111A-83	Sequence 83, Appl
47	26	86.7	63	6	US-08-478-029A-83	Sequence 83, Appl
48	26	86.7	107	3	US-08-961-166A-4	Patient No. 5198342
49	26	86.7	135	1	US-07-609-716-84	Sequence 84, Appl
50	26	86.7	135	1	US-07-609-716-84	Sequence 84, Appl
51	26	86.7	135	4	US-08-478-029A-84	Sequence 84, Appl
52	26	86.7	176	1	US-07-924-054-8	Sequence 8, Appl
53	26	86.7	176	1	US-08-076-011-2	Sequence 20, Appl
54	26	86.7	176	6	US-08-063-472B-20	Patient No. 5198342
55	26	86.7	210	1	US-08-234-783-2	Sequence 2, Appl
56	26	86.7	210	1	US-08-455-907-2	Sequence 2, Appl
57	26	86.7	210	1	US-08-455-907-2	Sequence 2, Appl
58	26	86.7	210	5	PCT-US95-0553-2	Sequence 2, Appl
59	26	86.7	211	3	US-08-433-263B-2	Sequence 2, Appl
60	26	86.7	343	2	US-09-019-216-3	Sequence 3, Appl
61	26	86.7	390	3	US-08-916-564A-2	Sequence 4, Appl
62	26	86.7	429	1	US-08-458-983-4	Sequence 4, Appl
63	26	86.7	429	5	PCT-US95-0553-4	Sequence 4, Appl
64	26	86.7	449	1	US-08-102-942A-6	Sequence 4, Appl
65	26	86.7	449	1	US-08-102-942A-6	Sequence 6, Appl
66	26	86.7	455	4	US-09-237-543-5	Sequence 15, Appl
67	26	86.7	557	1	US-08-313-288B-16	Sequence 16, Appl
68	26	86.7	560	2	US-08-329-492-5	Sequence 12, Appl
69	26	86.7	560	2	US-08-672-584-12	Sequence 12, Appl
70	26	86.7	568	1	US-08-331-615-3	Sequence 3, Appl
71	26	86.7	568	1	US-08-331-615-3	Sequence 3, Appl
72	26	86.7	1127	4	US-08-937-195-5	Sequence 3, Appl
73	26	86.7	1127	4	US-08-915-157-3	Sequence 3, Appl
74	26	86.7	1127	5	PCT-US96-97627-3	Sequence 3, Appl
75	26	86.7	1514	2	US-08-853-310-4	Sequence 4, Appl
76	26	86.7	1514	2	US-08-374-070C-2	Sequence 2, Appl
77	26	86.7	2126	3	US-08-963-188C-13	Sequence 13, Appl
78	26	86.7	2126	3	US-08-963-188C-13	Sequence 13, Appl
79	26	86.7	2987	2	US-08-407-562-29	Sequence 29, Appl
80	26	86.7	2987	2	US-08-407-562-29	Sequence 29, Appl
81	26	86.7	3959	2	US-08-970-269A-30	Sequence 30, Appl
82	26	86.7	3959	2	US-09-407-562-30	Sequence 30, Appl
83	25	83.3	106	3	US-07-636-551B-1	Sequence 1, Appl
84	25	83.3	106	3	US-08-963-168C-14	Sequence 14, Appl
85	25	83.3	111	3	US-08-963-168C-13	Sequence 15, Appl
86	25	83.3	111	3	US-08-963-168C-13	Sequence 15, Appl
87	25	83.3	126	3	US-08-963-168C-7	Sequence 7, Appl
88	25	83.3	126	3	US-08-963-168C-7	Sequence 8, Appl
89	25	83.3	132	3	US-08-963-168C-6	Sequence 6, Appl
90	25	83.3	132	3	US-08-963-168C-6	Sequence 6, Appl
91	25	83.3	136	3	US-08-963-168C-6	Sequence 6, Appl
92	25	83.3	197	4	US-08-975-215-4	Sequence 4, Appl
93	25	83.3	212	5	US-08-975-215-2	Sequence 4, Appl
94	25	83.3	212	5	US-08-975-215-2	Sequence 4, Appl
95	25	83.3	984	1	US-08-227-073-3	Sequence 120, Appl
96	25	83.3	984	1	US-08-184-009-120	Sequence 120, Appl
97	25	83.3	984	2	US-08-458-356-120	Sequence 120, Appl
98	25	83.3	984	4	US-08-460-736-120	Sequence 120, Appl
99	24	80.0	8	1	US-08-056-335-155	Sequence 155, Appl
100	24	80.0	8	1	US-08-056-335-155	Sequence 155, Appl

## ALIGNMENTS

RESULT 1  
US-08-826-712-11  
Sequence 11, Application US/08828712  
Patent No. 6015984  
GENERAL INFORMATION:  
APPLICANT: Schneck, Jonathan P.  
APPLICANT: O'Herrin, Sean  
TITLE OF INVENTION: Soluble Divalent and Multivalent  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Banner & Witcoff, Ltd.  
STREET: 1001 G Street, N.W.  
CITY: Washington  
STATE: D.C.  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: 08/826/712  
FILING DATE: 28-MAR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A. 32,141  
REGISTRATION/DOCKET NUMBER: 01107.73713  
TELEPHONE: (202)508-9100  
TELEFAX: (202)508-9299  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
HYPOTHETICAL: NO  
FRAGMENT TYPE: Internal  
ORIGINAL SOURCE: Internal  
INDIVIDUAL ISOLATE: Ig kappa peptide linker  
US-08-826-712-11

Query Match 100.0%; Score 30; DB 3; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLGGS 6  
DB 1 GSLGGS 6

RESULT 2  
US-09-063-276-11  
Sequence 11, Application US/09063276  
Patent No. 6015984  
GENERAL INFORMATION:  
APPLICANT: Schneck, Jonathan  
APPLICANT: O'Herrin, Sean  
TITLE OF INVENTION: Molecular Complexes which  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Banner & Witcoff  
US-09-063-276-11

STREET: 1001 G Street, NW  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTISO for Windows Version 2.0  
CURRENT APPLICATION DATA: US/09/063,276  
FILING DATE: 19-APRIL-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/828,712  
FILING DATE: 28-MAR-1997  
APPLICATION NUMBER: 60/014,367  
FILING DATE: 28-MAR-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Kagan, Sarah A. 32141  
REGISTRATION/DOCKET NUMBER: 01107.74154  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-508-9100  
TELEFAX: 202-508-9299  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 6 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-063-276-11

Query Match 100.0%; Score 30; DB 4; Length 6;  
Best Local Similarity 100.0%; Pred. No. 1.6e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLGGS 6  
DB 1 GSLGGS 6

RESULT 3  
US-08-635-121-2  
Sequence 2, Application US/08635121  
Patent No. 6015984  
GENERAL INFORMATION:  
APPLICANT: German, Irwin H.  
TITLE OF INVENTION: Tumor Suppressor Gene  
NUMBER OF SEQUENCES: 2  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Brumbaugh, Givens, Donohue & Raymond  
STREET: 30 Rockefeller Plaza  
STATE: NY  
COUNTRY: USA  
ZIP: 10112-0228  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: Version 1.5  
CURRENT APPLICATION DATA: US/08/635,121  
FILING DATE: 19-APRIL-1996  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:



NAME: Clark, Richard S  
 INFORMATION NUMBER: 26,154  
 REFERENCE/DOCKET NUMBER: A30588 - 165/33603  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 212-408-2558  
 TELEFAX: 212-765-2519  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1346 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 ANNOTATED FEATURES:  
 PHAGMOTYPE: NO  
 PHAGMOTYPE: Internal  
 ORIGINAL SOURCE:  
 US-08-635-121-2

Query Match 100.0% Score 30; DB 2; Length 1346;  
 Best Local Similarity 100.0%; Pred. No. 5,3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 960 GSIGS 965

RESULT 4  
 US-08-014-153D-45  
 Sequence 45, Application US/08014153D  
 GENERAL INFORMATION:  
 APPLICANT: Hadlock, Kenneth G.  
 Fong, Chin-Joo  
 TITLE OF INVENTION: Method and Assay for HTLV  
 NUMBER OF SEQUENCES: 45  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Genelabs Technologies, Inc.  
 STREET: 505 Penobscot Drive  
 CITY: Redwood City  
 STATE: CA, USA  
 COUNTRY: USA  
 ZIP: 94063  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentL Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/014,153D  
 FILING DATE: 05-Feb-1993  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/653,091  
 FILING DATE: 08-Feb-1991  
 APPLICATION NUMBER: 07/366,313  
 FILING DATE: 11-JUN-1989  
 APPLICATION NUMBER: US 06/948,270  
 FILING DATE: 31-DEC-1986  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Brookes, Allan A.  
 INFORMATION NUMBER: 36,373  
 REFERENCE/DOCKET NUMBER: G4C1P3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 369-9500  
 TELEFAX: (415) 368-0709  
 INFORMATION FOR SEQ ID NO: 45:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1346 amino acids  
 TYPE: amino acid

TOPOLOGY: linear  
 MOLECULE TYPE: amino acid  
 STRANDEDNESS: single  
 INFORMATION: SEQ ID NO: 45:  
 US-08-014-153D-45

Query Match 93.3% Score 28; DB 1; Length 21;  
 Best Local Similarity 83.3%; Pred. No. 16;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 10 GSIGS 15

RESULT 5  
 US-09-041-991A-8  
 Sequence 8, Application US/09041991A  
 Patent No. 6107278  
 GENERAL INFORMATION:  
 APPLICANT: Schnepf, H. Ernest  
 TITLE OF INVENTION: Peptides Active Against Pests  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Salliwanchik, Lloyd & Salliwanchik  
 STREET: 2421 N.W. 41st Street, Suite A-1  
 STATE: Gainesville  
 COUNTRY: USA  
 ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: IBM PC compatible  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/041,991A  
 FILING DATE: 13-MAR-1998  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Schnepf, H. Ernest  
 INFORMATION NUMBER: 39,355  
 REFERENCE/DOCKET NUMBER: NA-709  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (352) 375-8100  
 TELEFAX: (352) 372-5800  
 INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 633 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-041-991A-8

Query Match 93.3% Score 28; DB 3; Length 633;  
 Best Local Similarity 83.3%; Pred. No. 5,7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

DB 349 GSIGS 354

RESULT 6  
 US-09-041-991A-10  
 Sequence 10, Application US/09041991A  
 Patent No. 6107278  
 GENERAL INFORMATION:  
 APPLICANT: Schnepf, H. Ernest

APPLICANT: Narva, Kenneth E.  
 APPLICANT: Muller-Gohm, Judy  
 NUMBER OF SEQUENCES: 10  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Sallwanchik, Lloyd & Sallwanchik  
 STREET: 2421 N.W. 41st Street, Suite A-1  
 CITY: Gainesville  
 STATE: Florida  
 ZIP: 32606  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US-09/041,991A  
 FILING DATE: 13-MAR-1998  
 CLASSIFICATION: 514  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Sanders, Jay M., 39, 355  
 REGISTRATION NUMBER: 39, 355  
 TELEPHONE: (352) 375-8100  
 TELEFAX: (352) 375-8100  
 INFORMATION FOR SEQ ID NO: 10:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 633 amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-09-041-991A-10

Query Match 93.3% Score 28; DB 3; Length 633;  
 Best Local Similarity 83.3% Prod No. 5.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSIACS 6  
 DB 349 GSIACS 354

RESULT 7  
 US-08-704-931-6  
 Sequence 6, Application US/08704931  
 Patent No. 5885797  
 GENERAL INFORMATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Chen, Chao-Min (Amy)  
 REGISTRATION NUMBER: 38, 515  
 TELEPHONE: (206) 682-0446  
 TELEFAX: (206) 682-0446  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 163 amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-704-931-2

CLASSIFICATION: 536  
 ATTORNEY/AGENT INFORMATION:  
 NAME: King, Jeffrey J  
 REGISTRATION NUMBER: 38, 515  
 REFERENCE/DOCKET NUMBER: H011, P02  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 206-682-0446  
 TELEFAX: 206-682-0446  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 163 amino acids  
 type: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-704-931-6

Query Match 90.0% Score 27; DB 2; Length 163;  
 Best Local Similarity 83.3% Prod No. 2.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy 1 GSIACS 6  
 DB 120 GSIACS 125

RESULT 8  
 US-08-704-931-2  
 Sequence 2, Application US/08704931  
 Patent No. 5885797  
 GENERAL INFORMATION:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Chen, Chao-Min (Amy)  
 REGISTRATION NUMBER: 38, 515  
 TELEPHONE: (206) 682-0446  
 TELEFAX: (206) 682-0446  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 163 amino acids  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-704-931-2

Query Match 90.0% Score 27; DB 2; Length 246;  
 Best Local Similarity 83.3% Prod No. 3.3e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
DB 120 GALSOS 125

## RESULT 9

US-08-704-931-4  
Sequence 4, Application US/08704931  
Patent No. 5685797  
ORGANISM INFORMATION:  
APPLICANT: Chao-Min (AMY)  
APPLICANT: Kraut, No. 5685797oct  
APPLICANT: Groudnau, Mark  
APPLICANT: Melntrub, Harold  
TITLE OF INVENTION: No. 5685797el DNA Sequences Encoding Proteins  
TITLE OF INVENTION: INVOLVED IN MYOGENESIS  
CORRESPONDING ADDRESSES: 23  
ADDRESSER: Stratton Bullew, PLLC  
STREET: 1218 Third Avenue, Suite 1313  
CITY: Seattle  
STATE: WA  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
FILING DATE: NUMBER: US/08/704,931  
CLASSIFICATION: 536  
ATTORNEY/AGENT INFORMATION:  
NAME: King, Jeffrey J  
REGISTRATION NUMBER: 38,515  
REFERENCE/DOCKET NUMBER: H011.P02  
TELEPHONE: 206-683-1496  
TELEFAX: 206-682-0446  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 251 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-704-931-4

Query Match 90.0% Score 27; DB 2; Length 251;  
Best Local Similarity 83.3% Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
DB 120 GALSOS 125

RESULT 10  
US-09-237-543-4  
Sequence 4, Application US/09237543A  
Patent No. 6143540  
GENERAL INFORMATION:  
APPLICANT: Kapeller, Rosana  
TITLE OF INVENTION: NOVEL MOLECULES OF THE HK1D-1-RELATED PROTEIN FAMILY  
FILE REFERENCE: 035600/75631  
CURRENT APPLICATION NUMBER: US/09/237-543A  
CURRENT FILING DATE: 1999-01-25  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: Patentin Ver. 2.0  
SEQ ID NO 4

LENGTH: 254  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: eukaryotic  
US-09-237-543-4

Query Match 90.0% Score 27; DB 4; Length 254;  
Best Local Similarity 83.3% Pred. No. 3.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
DB 41 GSLOGS 46

RESULT 11  
US-08-633-148-4  
Sequence 4, Application US/08633148  
Patent No. 5864018  
ORGANISM INFORMATION:  
APPLICANT: MORSE, MICHAEL J.  
APPLICANT: NAGASHIMA, MARIO  
APPLICANT: HOLLANDER, DORIS A.  
TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION  
TITLE OF INVENTION: ENO-PRODUCT RECEPTOR POLYPEPTIDES AND USES THEREFOR  
NUMBER OF SEQUENCES: 23  
CORRESPONDING ADDRESSES:  
ADDRESSER: TOWNSEND & TOWNSEND  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
FILING DATE: NUMBER: US/08/633,148  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY ESO., MATTHEW B.  
REGISTRATION NUMBER: 39,787  
REFERENCE/DOCKET NUMBER: 014518-005600US  
TELEPHONE: 415-3924200  
TELEFAX: 415-3924200  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 318 amino acids  
TYPE: amino acid  
STANDARDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-633-148-4

Query Match 90.0% Score 27; DB 2; Length 318;  
Best Local Similarity 83.3% Pred. No. 4.3e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
DB 309 GSLOGS 314

RESULT 12  
US-09-237-543-2

Sequence 2, Application US/09237543A  
Patent No. 6143540  
GENERAL INFORMATION:  
APPLICANT: Kapeller, Rosana  
TITLE OF INVENTION: NOVEL MOLECULES OF THE HK1D-1-RELATED PROTEIN FAMILY  
FILE REFERENCE: US/09237543A  
CURRENT FILING DATE: 1999-01-26  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 2  
LENGTH: 326  
TYPE: PPT  
MOLECULE TYPE: Homo sapiens  
US-09-633-148-2

Query Match 90.0%; Score 27; DB 4; Length 326;  
Best Local Similarity 83.3%; Pred. No. 4,4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLOGS 6  
|||||  
DB 80 GSLOGS 85

RESULT 13  
US-08-633-148-2  
Sequence 4, Application US/08633148  
Patent No. 5864018  
GENERAL INFORMATION:  
APPLICANT: MORSE, MICHAEL J.  
APPLICANT: NAGASHIMA, MARIKO  
TITLE OF INVENTION: ANTIBODIES TO ADVANCED GLYCOSYLATION  
FILE REFERENCE: US/08633148  
CURRENT FILING DATE: 1999-01-26  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND & TOWNSEND & CREW LLP  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
APPLICATION NUMBER: US/08/633,148  
FILING DATE: 16-APR-1996  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: MURPHY ESO, MATTHEW B.  
REGISTRATION NUMBER: 39,787  
ADDRESS: 1000 LINCOLN STREET, 35TH FLOOR  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 326-2422  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
STANDARDNESS: acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-633-148-2

Query Match 90.0%; Score 27; DB 2; Length 340;  
Best Local Similarity 83.3%; Pred. No. 4,6e+02;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLOGS 6  
|||||  
DB 331 GSLOGS 336

RESULT 14  
US-08-630-172-6  
Sequence 6, Application US/08630172  
Patent No. 6060054  
GENERAL INFORMATION:  
APPLICANT: STATION, Uge  
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, 35th Floor  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.  
ZIP: 80203

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/630,172  
FILING DATE:  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: COLOMBO, JAMES J.  
REGISTRATION NUMBER: 32,020  
REFERENCE/DOCKET NUMBER: 2879-36  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (303) 863-9700  
TELEFAX: (303) 863-0223  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 340 amino acids  
STANDARDNESS: acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-630-172-6

Query Match 90.0%; Score 27; DB 3; Length 349;  
Best Local Similarity 83.3%; Pred. No. 4,7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSLOGS 6  
|||||  
DB 156 GSLOGS 161

RESULT 15  
US-09-375-419-6  
Sequence 6, Application US/09375419  
Patent No. 6264950  
GENERAL INFORMATION:  
APPLICANT: Uge  
TITLE OF INVENTION: NOVEL PRODUCT AND PROCESS FOR T  
NUMBER OF SEQUENCES: 41  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Sheridan Ross & McIntosh  
STREET: 1700 Lincoln Street, 35th Floor  
CITY: Denver  
STATE: Colorado  
COUNTRY: U.S.  
ZIP: 80203

COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/375,419  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/630,172  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Connell, Gary J.  
 REGISTRATION NUMBER: 32,020  
 REFERENCE/DOCKET NUMBER: 2879-36  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (303) 863-9900  
 TELEFAX: (303) 863-9223  
 INFORMATION FOR SEQ ID NO:  
 SEQUENCE CHARACTERISTICS: 6  
 LENGTH: 349 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-09-375-419-6

Query Match 90.0%; Score 27; DB 4; Length 349;  
 Best Local Similarity 83.3%; Pred. No. 4.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGSS 6  
 Db 156 GSIGCT 161

RESULT 16  
 US-08-464-266-2  
 Sequence 2, Application US/08464266  
 Patent No. 5616552  
 GENERAL INFORMATION:  
 APPLICANT: ORO, Ph.D., ANTHONY E.  
 APPLICANT: EVANS, Ph.D., RONALD M.  
 TITLE OF INVENTION: INSECT RETINOID-LIKE RECEPTOR  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS  
 NUMBER OF SEQUENCES: 5  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/464,266  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/013,975  
 FILING DATE: 04-FEB-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/497,935  
 FILING DATE: 22-FEB-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Retter, Stephen E.  
 REGISTRATION NUMBER: 31,192  
 REFERENCE/DOCKET NUMBER: P41 9966

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-546-4737  
 TELEFAX: 619-546-9392  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 513 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-464-266-2

Query Match 90.0%; Score 27; DB 1; Length 513;  
 Best Local Similarity 83.3%; Pred. No. 7.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGSS 6  
 Db 200 GSIGGS 205

RESULT 17  
 US-08-464-272-2  
 Sequence 2, Application US/08464272  
 Patent No. 5686691  
 GENERAL INFORMATION:  
 APPLICANT: ORO, Ph.D., ANTHONY E.  
 APPLICANT: EVANS, Ph.D., RONALD M.  
 TITLE OF INVENTION: INSECT RETINOID-LIKE RECEPTOR  
 TITLE OF INVENTION: COMPOSITIONS AND METHODS  
 NUMBER OF SEQUENCES:  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Pretty, Schroeder, Brueggemann & Clark  
 STREET: 444 South Flower Street, Suite 2000  
 CITY: Los Angeles  
 STATE: CA  
 COUNTRY: United States  
 ZIP: 90071  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/464,272  
 FILING DATE:  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/013,975  
 FILING DATE: 04-FEB-1993  
 APPLICATION NUMBER: US 07/497,935  
 FILING DATE: 22-FEB-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Retter, Stephen E.  
 REGISTRATION NUMBER: 31,192  
 REFERENCE/DOCKET NUMBER: P41 9950  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 619-546-4737  
 TELEFAX: 619-546-9392  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 513 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-464-272-2

Query Match 90.0%; Score 27; DB 1; Length 513;  
 Best Local Similarity 83.3%; Pred. No. 7.1e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSIGSS 6

Db 200 GSVCS 205

11:111

RESULT 18  
US-08-464-514-2  
Sequence 22 Application US/08464514  
Patent No. 6265177  
GENERAL INFORMATION:  
APPLICANT: EVANS, RONALD M.  
APPLICANT: MCKEON, MICHAEL B.  
APPLICANT: ORO, ANTHONY E.  
APPLICANT: SEARS, WILLIAM A.  
APPLICANT: WATSON, JAMES D.  
TITLE OF INVENTION: MULTIMERIC FORMS OF MEMBERS OF THE  
TITLE OF INVENTION: STEROID/THYROID SUPERFAMILY OF RECEPTORS WITH THE  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
STREET: 144 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: California  
COUNTRY: United States  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,514  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: 07/907,908  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: P41 9321  
REFERENCE/DOCKET NUMBER: P41 9321  
TELEPHONE: (619) 546-9392  
TELEFAX: (619) 546-9392  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-464-514-2

Query Match 90.0% Score 27; DB 4; Length 513;  
Best Local Similarity 83.3% Pred. No. 7.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSVCS 6  
11:111

Db 200 GSVCS 205

RESULT 19  
US-08-464-403-2  
Sequence 2 Application US/0846403  
Patent No. 628130  
GENERAL INFORMATION:  
APPLICANT: EVANS, RONALD M.  
APPLICANT: MCKEON, MICHAEL B.  
APPLICANT: ORO, ANTHONY E.  
APPLICANT: SEARS, WILLIAM A.  
APPLICANT: WATSON, JAMES D.  
TITLE OF INVENTION: MULTIMERIC FORMS OF MEMBERS OF THE

TITLE OF INVENTION: STEROID/THYROID SUPERFAMILY OF RECEPTORS WITH THE  
TITLE OF INVENTION: ULTRASPIRACLE RECEPTOR  
NUMBER OF SEQUENCES: 29  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PRETTY, SCHROEDER, BRUEGGEMANN & CLARK  
STREET: 144 South Flower Street, Suite 2000  
CITY: Los Angeles  
STATE: California  
COUNTRY: United States  
ZIP: 90071  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,403  
FILING DATE:  
CLASSIFICATION: 435  
PRIORITY APPLICATION DATA: US 07/907,908  
APPLICATION NUMBER: 07-JUL-1992  
FILING DATE: 02-JUL-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Reiter, Stephen E.  
REGISTRATION NUMBER: 31192  
REFERENCE/DOCKET NUMBER: P41 9321  
TELEPHONE: (619) 546-9392  
TELEFAX: (619) 546-9392  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 513 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-466-403-2

Query Match 90.0% Score 27; DB 4; Length 513;  
Best Local Similarity 83.3% Pred. No. 7.1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GSVCS 6  
11:111

Db 200 GSVCS 205

RESULT 20  
US-08-642-406A-22  
Sequence 22 Application US/08642406A  
Patent No. 5959177  
GENERAL INFORMATION:  
APPLICANT: Hehn, Mich B.  
APPLICANT: Helt, Andrew C.  
APPLICANT: INVER, Julian K.  
TITLE OF INVENTION: TRANSGENIC PLANTS EXPRESSING ASSEMBLED  
TITLE OF INVENTION: SECRETORY ANTIBODIES  
NUMBER OF SEQUENCES: 26  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: THE SCRIPPS RESEARCH INSTITUTE  
STREET: 10666 NO. 5959177th Torrey Pines Road, TPC-8  
CITY: La Jolla  
STATE: California  
COUNTRY: US  
ZIP: 92037  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/642,406A  
FILING DATE: 03-MAY-1996

CLASSIFICATION: 800  
PRIOR APPLICATION DATA: US 07/591,833  
FILING DATE: 27-OCT-1989  
APPLICATION NUMBER: US 07/427,765  
FILING DATE: 27-OCT-1989  
ATTORNEY/AGENT INFORMATION:  
NAME: Logan, April C.  
REGISTRATION NUMBER: 33,950  
RESIDENCE/DOCKET NUMBER: 184.2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 554-2937  
TELEFAX: (619) 554-6312  
INFORMATION FOR SEQ ID NO: 22:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 624 amino acids  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-642-406A-22

Query Match 90.0%; Score 27; DB 2; Length 624;  
Best Local Similarity 83.3%; Pred. No. 8.7e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
1 GSIAGS 6  
DB 250 GMLAGS 255

RESULT 21  
US-08-434-000A-2  
Sequence 2; Application US/08434000A  
Patent No. 6046037  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN K.-C. MA, THOMAS LEHNER  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
NUMBER OF SOURCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
ZIP: 90071 U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: MORD Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/434,000A  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: including application  
PRIOR APPLICATION DATA: described below:  
APPLICATION NUMBER: US/08/367,395  
FILING DATE: 12/30/94  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 552-0150  
TELEFAX: (619) 552-0159  
TELEX: 67-3510

TELEX: SEQUENCE LISTING  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 773 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
DESCRIPTION: Rabbit polyimmunoglobulin receptor  
US-08-434-000A-2

Query Match 90.0%; Score 27; DB 3; Length 773;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
1 GSIAGS 6  
DB 250 GMLAGS 255

RESULT 22  
US-09-312-157-2  
Sequence 2; Application US/09312157  
Patent No. 6303341  
GENERAL INFORMATION:  
APPLICANT: ANDREW C. HIATT, JULIAN K.-C. MA, THOMAS LEHNER  
TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION  
NUMBER OF SOURCES: 3  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Lyon & Lyon  
STREET: 633 West Fifth Street  
CITY: Los Angeles  
STATE: California  
ZIP: 90071 U.S.A.  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 MB  
MEDIUM TYPE: storage  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: IBM P.C. DOS 5.0  
SOFTWARE: MORD Perfect 5.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/312,157  
FILING DATE: 14-May-1999  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/434,000  
FILING DATE: <unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Guise, Jeffrey W.  
REGISTRATION NUMBER: 34,613  
REFERENCE/DOCKET NUMBER: 212/127  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 552-8400  
TELEFAX: (619) 552-0159  
TELEX: 67-3510  
SEQUENCE LISTING  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 773 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
DESCRIPTION: Rabbit polyimmunoglobulin receptor  
US-09-312-157-2

Query Match 90.0%; Score 27; DB 4; Length 773;  
Best Local Similarity 83.3%; Pred. No. 1.1e+03;

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GSLGCS-6  
1:|||||

DB 250 GNLGCS 255

RESULT 23

US-08-680-326-41

Sequence 41: Application US/08680326

Patent No. 59257337

GENERAL INFORMATION:

ATTORNEY/AGENT INFORMATION:

APPLICANT: TIMOTHY M.

APPLICANT: BOSCH MARINX

APPLICANT: TODARO, GEORGE J.

TITLE OF INVENTION: DNA POLYMERASE OF GAMMA HERPES VIRUSES

TITLE OF INVENTION: ASSOCIATED WITH KAPOSI'S SARCOMA AND RETROPERITONEAL

NUMBER OF SEQUENCES: 152

NUMBER OF SEQUENCES: 152

ADDRESS: 755 Page Mill Road

STREET: 755 Page Mill Road

CITY: Palo Alto

STATE: California

COUNTRY: USA

ZIP: 94304-1018

COMPUTER: IBM PC compatible

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/680,326

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Schiff, J. Michael

REGISTRATION NUMBER: 40,253

REFERENCE/DOCKET NUMBER: 29938-20001.00

TELEPHONE: (415) 813-5800

TELEX: 706141

INFORMATION FOR SEQ ID NO: 41:

SEQUENCE CHARACTERISTICS:

LENGTH: 985 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-680-326-41

Query Match 90.0%: Score 27: DB 2: Length 985:

Best Local Similarity 83.3%: Pred. No. 1.4e+03:

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GNLGCS 6  
1:|||||

DB 133 GNLGCS 138

RESULT 24

US-08-667-941-13

Sequence 13: Application US/08667941

Patent No. 5977337

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M

APPLICANT: Du, Run-Pan

APPLICANT: Wang, Yan-Ping

APPLICANT: Klein, Michael H

TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA

NUMBER OF SEQUENCES: 67

CORRESPONDENCE ADDRESS:

ADDRESS: 6th Floor, 350 University Avenue

STREET: 6th Floor, 350 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/667,941

FILING DATE: 03-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael

REFERENCE/DOCKET NUMBER: 1038-601 MIS:jb

TELEPHONE: (416) 595-1155

TELEFAX: (416) 595-1163

INFORMATION FOR SEQ ID NO: 13:

SEQUENCE CHARACTERISTICS:

LENGTH: 985 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-08-667-941-13

Query Match 90.0%: Score 27: DB 2: Length 985:

Best Local Similarity 83.3%: Pred. No. 1.4e+03:

Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GNLGCS 6  
1:|||||

DB 175 GNLGCS 180

RESULT 25

US-08-667-941-17

Sequence 17: Application US/08667941

Patent No. 5977337

GENERAL INFORMATION:

APPLICANT: Loosmore, Sheena M

APPLICANT: Du, Run-Pan

APPLICANT: Wang, Yan-Ping

APPLICANT: Klein, Michael H

TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA

CORRESPONDENCE ADDRESS:

ADDRESS: 6th Floor, 350 University Avenue

STREET: 6th Floor, 350 University Avenue

CITY: Toronto

STATE: Ontario

COUNTRY: Canada

ZIP: M5G 1R7

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/667,941

FILING DATE: 03-JUN-1997

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Stewart, Michael

REGISTRATION NUMBER: 24,973



REFERENCE/DOCKET NUMBER: 1038-601 MIS:b  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1155  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 985 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-08-867-941-17

Query Match 90.0%; Score 27; DB 4; Length 985;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 175 GALSOS 180

RESULT 26  
 US-09-074-658-13  
 Sequence 12; Application US/09074658  
 Patent No. 6184371  
 GENERAL INFORMATION:  
 APPLICANT: Loosmore, Sheena M  
 APPLICANT: Run-Pan Du  
 APPLICANT: Quijun Wang  
 APPLICANT: Wang, Yan-Ping  
 APPLICANT: Kiehn, Michael H  
 TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
 NUMBER OF SEQUENCES: 78  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Slim & McBurney  
 STREET: 6th Floor, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 FILING DATE: 08-MAY-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-795  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1155  
 INFORMATION FOR SEQ ID NO: 13:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 985 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-074-658-13

Query Match 90.0%; Score 27; DB 4; Length 985;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 175 GALSOS 180

RESULT 27  
 US-09-074-658-17  
 Sequence 17; Application US/09074658  
 Patent No. 6184371  
 GENERAL INFORMATION:  
 APPLICANT: Loosmore, Sheena M  
 APPLICANT: Run-Pan Du  
 APPLICANT: Quijun Wang  
 APPLICANT: Wang, Yan-Ping  
 APPLICANT: Kiehn, Michael H  
 TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
 NUMBER OF SEQUENCES: 78  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Slim & McBurney  
 STREET: 6th Floor, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: M5G 1R7  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 FILING DATE: 08-MAY-1998  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Stewart, Michael I  
 REGISTRATION NUMBER: 24,973  
 REFERENCE/DOCKET NUMBER: 1038-795  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (416) 595-1155  
 TELEFAX: (416) 595-1155  
 INFORMATION FOR SEQ ID NO: 17:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 985 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 US-09-074-658-17

Query Match 90.0%; Score 27; DB 4; Length 985;  
 Best Local Similarity 83.3%; Pred. No. 1.4e+03;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GSLOGS 6  
 DB 175 GALSOS 180

RESULT 28  
 US-08-867941-12  
 Sequence 12; Application US/08867941  
 Patent No. 5977337  
 GENERAL INFORMATION:  
 APPLICANT: Loosmore, Sheena M  
 APPLICANT: Du, Run-Pan  
 APPLICANT: Wang, Quijun  
 APPLICANT: Wang, Yan-Ping  
 APPLICANT: Kiehn, Michael H  
 TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
 NUMBER OF SEQUENCES: 67  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Slim & McBurney  
 STREET: 6th Floor, 330 University Avenue  
 CITY: Toronto  
 STATE: Ontario  
 COUNTRY: Canada

```

? ZIP: MSG 187
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/867,941
? FILING DATE: 03-JUN-1997
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Stewart, Michael I
? REGISTRATION NUMBER: 973
? RESC/DOCKET NUMBER: 1038-681.MIS:jb
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1000 amino acids
? STRANDEDNESS: single
? TOPOLOGY: linear
?
US-08-867-941.12

```

```

Query Match          90.0%; Score 27; DB 2; Length 1000;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSIQGS 6
DB 190 GAIQGS 195

```

```

RESULT 29
US-08-867-941-16
? Sequence 16, Application US/08867941
? Patent No. 5977337
? GENERAL INFORMATION:
? APPLICANT: Loosmore, Sheena M
? APPLICANT: Du, Kun-Pan
? APPLICANT: Yang, Yan-Ping
? APPLICANT: Yang, Yan-Ping
? APPLICANT: Klein, Michael H
? TITLE OF INVENTION: LACTOPERRIN RECEPTOR GENES OF MORAXELLA
? NUMBER OF SEQUENCES: 67
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: SIm & McBurney
? STREET: 6th Floor, 330 University Avenue
? CITY: Toronto
? STATE: Ontario
? COUNTRY: Canada
?
ZIP: MSG 187
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/867,941
? FILING DATE: 03-JUN-1997
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Stewart, Michael I
? REGISTRATION NUMBER: 973
? RESC/DOCKET NUMBER: 1038-681.MIS:jb
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 16:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1000 amino acids

```

```

? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
US-08-867-941-16

```

```

Query Match          90.0%; Score 27; DB 2; Length 1000;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSIQGS 6
DB 190 GAIQGS 195

```

```

RESULT 30
US-09-074-658-12
? Sequence 12, Application US/09074658
? Patent No. 658137
? GENERAL INFORMATION:
? APPLICANT: Loosmore, Sheena M
? APPLICANT: Run-Pan Du
? APPLICANT: Qun-Jun Wang
? APPLICANT: Yang, Yan-Ping
? APPLICANT: Klein, Michael H
? TITLE OF INVENTION: LACTOPERRIN RECEPTOR GENES OF MORAXELLA
? NUMBER OF SEQUENCES: 78
? CORRESPONDENCE ADDRESSES:
? ADDRESSEE: SIm & McBurney
? STREET: 6th Floor, 330 University Avenue
? CITY: Toronto
? STATE: Ontario
? COUNTRY: Canada
?
ZIP: MSG 187
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/074,658
? FILING DATE: 08-MAY-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? NAME: Stewart, Michael I
? REGISTRATION NUMBER: 24,973
? REFERENCE/DOCKET NUMBER: 1038-795
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 12:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1000 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
?
US-09-074-658-12

```

```

Query Match          90.0%; Score 27; DB 4; Length 1000;
Best Local Similarity 83.3%; Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 1 GSIQGS 6
DB 190 GAIQGS 195

```

```

RESULT 31
US-09-074-658-16
? Sequence 16, Application US/09074658
? Patent No. 658137
? GENERAL INFORMATION:

```

```

? APPLICANT: Loosmore, Sheena M
? APPLICANT: Run-Pan Du
? APPLICANT: Quljun Wang
? APPLICANT: Quljun Wang
? APPLICANT: Kiehl, Ian Ping
? APPLICANT: Kiehl, Ian Ping
? TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
? NUMBER OF SEQUENCES: 78
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Slim & McBurney
? STREET: 6th Floor, 330 University Avenue
? CITY: Toronto
? STATE: Ontario
? COUNTRY: Canada
? ZIP: M5G 1R7
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? CURRENT APPLICATION DATA: PC-DOS/MS-DOS
? SOFTWARE: IBM PC compatible
? APPLICATION NUMBER: US/09/074,658
? FILING DATE: 08-MAY-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 1073
? REFERENCE/DOCKET NUMBER: 1038-795
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 16:
? SOURCE: CHARACTERISTICS:
? SUBJECT: amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-09-074-658-16

Query Match          90.0%: Score 27; DB 4; Length 1000;
Best Local Similarity 83.3%: Pred. No. 1.4e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSLGGS 6
DB      190 GALSGS 195

RESULT 32
US-08-484-438-9
? Sequence 9, Application US/08484438
? Patent No. 5811098
? GENERAL INFORMATION:
? APPLICANT: Loosmore, Sheena M
? APPLICANT: Run-Pan Du
? APPLICANT: Quljun Wang
? APPLICANT: Quljun Wang
? APPLICANT: Kiehl, Ian Ping
? APPLICANT: Kiehl, Ian Ping
? TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
? NUMBER OF SEQUENCES: 78
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Slim & McBurney
? STREET: 6th Floor, 330 University Avenue
? CITY: Toronto
? STATE: Ontario
? COUNTRY: Canada
? ZIP: M5G 1R7
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? CURRENT APPLICATION DATA: PC-DOS/MS-DOS
? SOFTWARE: IBM PC compatible
? APPLICATION NUMBER: US/09/074,658
? FILING DATE: 08-MAY-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 1073
? REFERENCE/DOCKET NUMBER: 1038-795
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 16:
? SOURCE: CHARACTERISTICS:
? SUBJECT: amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-09-074-658-16

```

```

? SOFTWARE: PatentID Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/484,438
? FILING DATE: 07-JUN-1995
? CLASSIFICATION: 530
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: 08/23,442
? FILING DATE: 08-MAY-1998
? APPLICATION NUMBER: US 08/150,704
? FILING DATE: 10-NOV-1993
? CLASSIFICATION: 530
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: US 07/981,165
? FILING DATE: 24-NOV-1992
? CLASSIFICATION: 530
? ATTORNEY/AGENT INFORMATION:
? NAME: Mistrock, S. Leslie
? REGISTRATION NUMBER: 18,872
? REFERENCE/DOCKET NUMBER: 5624-230
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 790-9090
? INFORMATION FOR SEQ ID NO: 9:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1342 amino acids
? TYPE: amino acid
? STRANDEDNESS: unknown
? TOPOLOGY: unknown
? MOLECULE TYPE: protein
? US-08-484-438-9

Query Match          90.0%: Score 27; DB 2; Length 1342;
Best Local Similarity 83.3%: Pred. No. 1.9e+03;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 GSLGGS 6
DB      1060 GNLGGS 1065

RESULT 33
US-09-074-658-15
? Sequence 15, Application US/09074658
? Patent No. 6184371
? GENERAL INFORMATION:
? APPLICANT: Loosmore, Sheena M
? APPLICANT: Run-Pan Du
? APPLICANT: Quljun Wang
? APPLICANT: Quljun Wang
? APPLICANT: Kiehl, Ian Ping
? APPLICANT: Kiehl, Ian Ping
? TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA
? NUMBER OF SEQUENCES: 78
? CORRESPONDENCE ADDRESS:
? ADDRESSER: Slim & McBurney
? STREET: 6th Floor, 330 University Avenue
? CITY: Toronto
? STATE: Ontario
? COUNTRY: Canada
? ZIP: M5G 1R7
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? CURRENT APPLICATION DATA: PC-DOS/MS-DOS
? SOFTWARE: IBM PC compatible
? APPLICATION NUMBER: US/09/074,658
? FILING DATE: 08-MAY-1998
? CLASSIFICATION: 435
? ATTORNEY/AGENT INFORMATION:
? REGISTRATION NUMBER: 1073
? REFERENCE/DOCKET NUMBER: 1038-795
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (416) 595-1155
? TELEFAX: (416) 595-1163
? INFORMATION FOR SEQ ID NO: 16:
? SOURCE: CHARACTERISTICS:
? SUBJECT: amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? US-09-074-658-16

```

REFERENCE/DOCKET NUMBER: 1038-795  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INVENTOR: (US) 1038-795  
SEQUENCE CHARACTERISTICS: 15:  
LENGTH: 2432 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-074-658-15

Query Match 90.0%: Score 27; DB 4; Length 2432;  
Best Local Similarity 83.3%; Pred. No. 3,6e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSFGS 6  
DB 1084 GATGGS 1089

RESULT 34  
US-09-074-658-11  
Sequence 11, Application US/090746558  
Patent No. 6184371  
GENERAL INFORMATION:  
INVENTOR: Ron Sheena M  
APPLICANT: Ron Sheena M  
APPLICANT: Qutun Wang  
APPLICANT: Yang, Yan-Ping  
TITLE OF INVENTION: LACTOFERRIN RECEPTOR GENES OF MORAXELLA  
NUMBER OF SEQUENCES: 78  
CORRESPONDENCE ADDRESS:  
STREET: 6th floor, 330 University Avenue  
CITY: Toronto  
STATE: Ontario  
COUNTRY: Canada  
ZIP: M5G 1R7  
COMPUTER READABLE FORM:  
FILE TYPE: PC compatible  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/074,658  
FILING DATE: 08-MAY-1998  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Stewart, Michael I  
REGISTRATION NUMBER: 24,973  
REFERENCE/DOCKET NUMBER: 1038-795  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (416) 595-1155  
TELEFAX: (416) 595-1163  
INVENTOR: (US) 1038-795  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2439 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-09-074-658-11

Query Match 90.0%: Score 27; DB 4; Length 2439;  
Best Local Similarity 83.3%; Pred. No. 3,6e+03;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GSFGS 6  
DB 1568 GATGGS 1093

RESULT 35  
US-07-609-716-16  
Sequence 16, Application US/07609716  
Patent No. 5514581  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A.  
APPLICANT: Cappello, Joseph  
TITLE OF INVENTION: Functional Recombinantly Prepared  
TITLE OF INVENTION: Synthetic Protein Polymer  
NUMBER OF SEQUENCES: 118  
CORRESPONDENCE ADDRESS:  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
FILE TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/609,716  
FILING DATE: 06-NOV-1990  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Rowland, Berttram I  
REGISTRATION NUMBER: 20015  
REFERENCE/DOCKET NUMBER: A-55186-3/BIR  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-381-7469  
TELEFAX: 415-381-7469  
INFORMATION FOR SEQ. ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MODULE TYPE: peptide  
US-07-609-716-16

Query Match 86.7%: Score 26; DB 1; Length 15;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSFGS 6  
DB 4 GSFGS 9

RESULT 36  
US-07-609-716-67  
Sequence 6, Application US/07609716  
Patent No. 5514581  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A.  
APPLICANT: Cappello, Joseph  
TITLE OF INVENTION: Functional Recombinantly Prepared  
TITLE OF INVENTION: Synthetic Protein Polymer  
NUMBER OF SEQUENCES: 118  
CORRESPONDENCE ADDRESS:  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111  
COMPUTER READABLE FORM:  
FILE TYPE: Floppy disk

COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA  
 APPLICATION NUMBER: 08/07/609,716  
 FILING DATE: 06-NOV-1990  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Rowland, Bertam I  
 REGISTRATION NUMBER: 20015  
 TELEPHONE: 415-398-3249  
 TELEFAX: 415-398-3249  
 INFORMATION FOR SEQ ID NO: 67:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-07-609-716-67

Query Match 86.7%; Score 26; DB 1; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSFGS 6  
 |||||  
 DB 4 GSFGS 9

RESULT 37  
 US-08-477-5098-37  
 Sequence 37, Application US/084775098  
 Patent No. 5770697  
 GENERAL INFORMATION:  
 APPLICANT: Ferrari, Franco A  
 APPLICANT: Cappelletto, Joseph  
 APPLICANT: Cappelletto, John W  
 APPLICANT: Dorfman, Mary A  
 TITLE OF INVENTION: No. 5770697el Peptides Comprising Repetitive  
 TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same  
 NUMBER OF SEQUENCES: 112  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fleury, Hobach, Test, Albritton & Herbert  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: US  
 ZIP: 94111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/477,5098  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 PRIOR APPLICATION NUMBER: US 08/175,155  
 FILING DATE: 29-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/053,049  
 FILING DATE: 22-APR-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/114,618  
 FILING DATE: 22-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/927,258  
 FILING DATE: 04-NOV-1986

ATTORNEY/AGENT INFORMATION:  
 NAME: Treccarlin, Richard F.  
 REGISTRATION NUMBER: 31,801  
 REFERENCE/DOCKET NUMBER: A-55186-7/RT/MTK  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-781-1989  
 TELEFAX: 415-398-3249  
 INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-477-5098-37

Query Match 86.7%; Score 26; DB 1; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 27;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSFGS 6  
 |||||  
 DB 4 GSFGS 9

RESULT 38  
 US-08-482-0858-37  
 Sequence 37, Application US/084820858  
 Patent No. 6018040  
 GENERAL INFORMATION:  
 APPLICANT: Ferrari, Franco A.  
 APPLICANT: Richardson, Charles  
 APPLICANT: Chambers, James  
 APPLICANT: Causey, Stuart  
 APPLICANT: Collock, Thomas J.  
 APPLICANT: Cappelletto, Joseph  
 APPLICANT: Cappelletto, John W  
 APPLICANT: Cappelletto, John W  
 TITLE OF INVENTION: No. 6018040el Peptides Comprising Repetitive  
 TITLE OF INVENTION: Units of Amino Acids and DNA Sequences Encoding the Same  
 NUMBER OF SEQUENCES: 112  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fleury, Hobach, Test, Albritton & Herbert  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: US  
 ZIP: 94111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/482,0858  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 PRIOR APPLICATION NUMBER: US 06/927,258  
 FILING DATE: 04-NOV-1986  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/114,618  
 FILING DATE: 29-OCT-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/053,049  
 FILING DATE: 22-APR-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/175,155  
 FILING DATE: 29-DEC-1993  
 PRIOR APPLICATION DATA:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Treccarlin, Richard F.  
 REGISTRATION NUMBER: 31,801  
 REFERENCE/DOCKET NUMBER: A-55186-6/RT/MTK

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
KEY: amino acids  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-482-0856-37

Query Match 86.7% Score 26; DB 3; Length 15;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSIGS 6  
DB 1 1 1 1  
4 GSFGS 9

RESULT 39  
US-08-475-411A-16  
Sequence 16, Application US/08475411A  
Patent No. 6140072  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A.  
INVENTOR: Ferrari, Franco A.  
TITLE OF INVENTION: Functional Protein Prepared  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,411A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIORITY NUMBER: US 07/609,716  
FILING DATE: 06-NOV-1990  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/269,429  
FILING DATE: 09-NOV-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/114,618  
FILING DATE: 29-OCT-1987  
PERSONAL INFORMATION:  
APPLICATION NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Treacatan, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-9/RET/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 16:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
KEY: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-08-475-411A-16

Query Match 86.7% Score 26; DB 4; Length 15;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSIGS 6  
DB 1 1 1 1  
4 GSFGS 9

RESULT 40  
US-08-475-411A-67  
Sequence 67, Application US/08475411A  
Patent No. 6140072  
GENERAL INFORMATION:  
APPLICANT: Ferrari, Franco A.  
INVENTOR: Ferrari, Franco A.  
TITLE OF INVENTION: Functional Protein Prepared  
NUMBER OF SEQUENCES: 119  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Flehr, Hohbach, Test, Albritton & Herbert  
STREET: Four Embarcadero Center, Suite 3400  
CITY: San Francisco  
STATE: CA  
COUNTRY: US  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent In Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/475,411A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
PRIORITY NUMBER: US 07/609,716  
FILING DATE: 06-NOV-1990  
APPLICATION DATA:  
APPLICATION NUMBER: US 07/269,429  
FILING DATE: 09-NOV-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/114,618  
FILING DATE: 29-OCT-1987  
PRIOR APPLICATION DATA:  
PRIORITY NUMBER: US 06/927,258  
FILING DATE: 04-NOV-1986  
ATTORNEY/AGENT INFORMATION:  
NAME: Treacatan, Richard F.  
REGISTRATION NUMBER: 31,801  
REFERENCE/DOCKET NUMBER: A-55186-9/RET/MTK  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-781-1989  
TELEFAX: 415-398-3249  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 15 amino acids  
KEY: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-475-411A-67

Query Match 86.7% Score 26; DB 4; Length 15;  
Best Local Similarity 83.3%; Pred. No. 27;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GSIGS 6

Mon Feb 4 08:21:05 2002

DB 4 GSFCGS 9

Search completed: February 4, 2002, 08:01:38  
Job time: 82 sec

**THIS PAGE BLANK (USPTO)**





## ALIGNMENTS

RESULT 1  
US-08-828-712-10

Sequence 10, Application US/08828712

Patent No. 6015884

GENERAL INFORMATION:

APPLICANT: Schneck, Jonathan P.

APPLICANT: O'Heirin, Sean

TITLE OF INVENTION: Soluble Divalent and Multivalent

TITLE OF INVENTION: Soluble Divalent and Multivalent

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff, Ltd.

STREET: 1001 G Street, N.W.

CITY: Washington

STATE: D.C.

COUNTRY: U.S.A.

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent release #1.0, Version #1.30

APPLICATION NUMBER: US/08/828,712

FILING DATE: 28-MAR-1997

CLASSIFICATION: 514

ATTORNEY/AGENT INFORMATION:

NAME: Kagan, Sarah A.

REGISTRATION NUMBER: 32,141

TELEPHONE: (202)508-9100

TELEFAX: (202)508-9100

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: linear

MOLECULE TYPE: peptide

HYPOTHEICAL: NO

FRAGMENT TYPE: Internal

ORIGINAL SOURCE:

INDIVIDUAL ISOLATE: 19c1 peptide linker

US-08-828-712-10

Query Match

Best Local Similarity 100.0%; Score 33; DB 3; Length 6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTSG 6

Db 1 GGGTSG 6

RESULT 2

US-09-063-276-10

Sequence 10, Application US/09063276

Patent No. 6015884

GENERAL INFORMATION:

APPLICANT: Schneck, Jonathan

APPLICANT: O'Heirin, Sean

TITLE OF INVENTION: Molecular Complexes Which

TITLE OF INVENTION: Modify Immune Responses

NUMBER OF SEQUENCES: 20

CORRESPONDENCE ADDRESS:

ADDRESSEE: Banner & Witcoff

STREET: 1001 G Street, NW

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20001

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: PASTISO for Windows Version 2.0

APPLICATION NUMBER: US/09/063,276

FILING DATE: 21-APR-1998

APPLICATION DATA:

APPLICATION NUMBER: 08/828,712

FILING DATE: 28-MAR-1997

APPLICATION NUMBER: 60/014,367

FILING DATE: 28-MAR-1996

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 32141

REFERENCE/DOCKET NUMBER: 01107,74154

TELEPHONE: 202-508-9100

TELEFAX: 202-508-9299

INFORMATION FOR SEQ ID NO: 10:

SEQUENCE CHARACTERISTICS:

LENGTH: 6 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

US-09-063-276-10

Query Match

Best Local Similarity 100.0%; Score 33; DB 4; Length 6;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTSG 6

Db 1 GGGTSG 6

RESULT 3

US-08-974-546-1

Sequence 1, Application US/08974546

Patent No. 5945287

GENERAL INFORMATION:

APPLICANT: Schneck, Jonathan L.

APPLICANT: Corley, Neil C.

TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS

TITLE OF INVENTION: TWO HUMAN HEAT SHOCK PROTEIN HOMOLOGS

NUMBER OF SEQUENCES: 6

CORRESPONDENCE ADDRESS:

ADDRESSEE: PharmaciaLabs, Inc.

STREET: 3174 Porter Dr.

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: DOS

SOFTWARE: PASTISO for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/974,546

FILING DATE: Filed Herein

Prior Application Number:

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: US, JUNE 5, 1995  
 REGISTRATION NUMBER: 6, 749  
 REFERENCE/DOCKET NUMBER: PP-0428  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 650-885-0555  
 TELEFAX: 650-845-4166  
 INFORMATION FOR SEQ ID NO. 1:  
 SEQUENCE CHARACTERISTICS:  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: BRAIUT21  
 CLONE: 2525691  
 US-08-974-546-1

Query Match 100.0%; Score 33; DB 2; Length 348;  
 Best Local Similarity 100.0%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 DB 75 GGGTSG 80

RESULT 4  
 US-09-415-522-6  
 Sequence 6, Application US/09415522A  
 Patent No. 6291660  
 GENERAL INFORMATION:  
 APPLICANT: Philippsen, Peter  
 APPLICANT: Wendland, Juergen  
 TITLE OF INVENTION: Development  
 TITLE OF INVENTION: Fungal Genes Required For No. 6291660mal Growth And  
 FILE REFERENCE: CGC2046  
 CURRENT APPLICATION NUMBER: US/09/415, 522A  
 FILING DATE: 1999-10-08  
 NUMBER OF SEQ NOS: 28, 2, 0  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 6  
 LENGTH: 2071  
 TYPE: PRT  
 ORGANISM: Ashbya gossypii  
 US-09-415-522-6

Query Match 100.0%; Score 33; DB 4; Length 2071;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 DB 199 GGGTSG 204

RESULT 5  
 US-08-461-9908-14  
 Sequence 14, Application US/084619908  
 Patent No. 5851810  
 GENERAL INFORMATION:  
 APPLICANT: JOHNS, BLANCHARD  
 TITLE OF INVENTION: NUCLEIC ACID ENCODING RHODOCOCCLUS  
 NUMBER OF SEQUENCES: 30  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: AMSTER, ROTHSTEIN & EBERSTEIN  
 STREET: 90 PARK AVENUE  
 CITY: NEW YORK

STATE: NEW YORK  
 COUNTRY: U.S.A.  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 FILING DATE: JUNE 5, 1995  
 OPERATING SYSTEM: IBM PC COMPATIBLE  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,9908  
 FILING DATE: JUNE 5, 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: US, JUNE 5, 1995  
 REGISTRATION NUMBER: 34, 287  
 REFERENCE/DOCKET NUMBER: 96700/370  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 697-5995  
 TELEFAX: (212) 286-0854 or 286-0082  
 INFORMATION FOR SEQ ID NO. 14:  
 SEQUENCE CHARACTERISTICS:  
 TYPE: AMINO ACID  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 DESCRIPTION: NO  
 HYPOTHEICAL: NO  
 IMMEDIATE SOURCE: STENOTHERMOPHILUS  
 ORGANISM: STENOTHERMOPHILUS  
 INDIVIDUAL ISOLATE: ALANINE DEHYDROGENASE  
 US-08-461-9908-14

Query Match 90.9%; Score 30; DB 2; Length 32;  
 Best Local Similarity 83.3%; Pred. No. 64;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSG 6  
 DB 6 GGGTSG 11

RESULT 6  
 US-08-461-9908-26  
 Sequence 26, Application US/084619908  
 Patent No. 5851810  
 GENERAL INFORMATION:  
 APPLICANT: JOHNS, BLANCHARD  
 TITLE OF INVENTION: NUCLEIC ACID ENCODING RHODOCOCCLUS  
 TITLE OF INVENTION: PHENYLALANINE DEHYDROGENASE  
 NUMBER OF SEQUENCES: 30  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: AMSTER, ROTHSTEIN & EBERSTEIN  
 STREET: 90 PARK AVENUE  
 CITY: NEW YORK  
 STATE: NEW YORK  
 COUNTRY: U.S.A.  
 ZIP: 10016  
 COMPUTER READABLE FORM:  
 FILING DATE: JUNE 5, 1995  
 OPERATING SYSTEM: IBM PC COMPATIBLE  
 SOFTWARE: ASCII  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,9908  
 FILING DATE: JUNE 5, 1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: US, JUNE 5, 1995  
 REGISTRATION NUMBER: 34, 287  
 REFERENCE/DOCKET NUMBER: 96700/370  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 697-5995  
 TELEFAX: (212) 286-0854 or 286-0082

TELEX: TXN 710-581-4766  
 INFORMATION: SEQ ID NO: 26;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 95  
 TYPE: AMINO ACID  
 TOPOLOGY: LINEAR  
 MOLECULE TYPE: PROTEIN  
 HYDROLYZABLE: NO  
 ORGANISM: B. STEAROTHERMOPHILUS  
 INDIVIDUAL ISOLATE: ALANINE DEHYDROGENASE  
 US-08-461-9908-26

Query Match 90.9% Score 30; DB 2; Length 95;  
 Best Local Similarity 83.3%; Pred. No. 1.7e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGCTSG 6  
 DB 67 GGCTAG 72

RESULT 7  
 US-09-028-934-29  
 Sequence 29, Application US/09028934  
 Patent No. 6117670  
 GENERAL INFORMATION:  
 APPLICANT: Ligon, James M.  
 APPLICANT: Hill, Dwight S.  
 APPLICANT: Hamner, Scott Joseph  
 APPLICANT: van Pee, Karl-Heinz  
 APPLICANT: Klinef, Sabine  
 APPLICANT: Young, Thomas R.  
 TITLE OF INVENTION: Pyrolytic Biosynthesis Genes and Uses  
 TITLE OF INVENTION: Thereof  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: No. 6117670atlis Corporation  
 STREET: 3054 Cornwallis Road  
 CITY: Research Triangle Park  
 STATE: NC  
 COUNTRY: USA  
 ZIP: 27709  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/028-934  
 FILING DATE: 08-JUN-1994  
 PRIORITY DATA:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/729,214  
 FILING DATE: 09-OCT-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/258,261  
 FILING DATE: 08-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Regus, J. Timothy  
 REGISTRATION NUMBER: 241  
 REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8587  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 29;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 537 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 US-09-028-934-29

Query Match 90.9% Score 30; DB 3; Length 537;  
 Best Local Similarity 83.3%; Pred. No. 7.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGCTSG 6  
 DB 12 GGCTAG 17

RESULT 8  
 US-08-258-261B-2  
 Sequence 2, Application US/08258261B  
 Patent No. 5639949  
 GENERAL INFORMATION:  
 APPLICANT: Schupp, Thomas  
 APPLICANT: Hamner, Scott Joseph  
 APPLICANT: Bock, James Joseph  
 APPLICANT: Hill, Dwight Steven  
 APPLICANT: Ryals, John Andrew  
 APPLICANT: Gaffney, Thomas Deane  
 APPLICANT: Lam, Stephen Ting  
 APPLICANT: Hammer, Phillip E.  
 APPLICANT: Klinef, Sabine  
 APPLICANT: Young, Thomas R.  
 TITLE OF INVENTION: The synthesis of  
 TITLE OF INVENTION: antipathogenic substances  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Ciba-Geigy Corporation  
 STREET: 7 Skyline Drive  
 CITY: Hawthorne  
 STATE: NJ  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/258-261B  
 FILING DATE: 08-JUN-1994  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/457,205  
 FILING DATE: 01-JUN-1995  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Blandt, Robert  
 REGISTRATION NUMBER: 35,129  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 2;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 538 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-258-261B-2

Query Match 90.9% Score 30; DB 1; Length 538;  
 Best Local Similarity 83.3%; Pred. No. 7.9e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGCTSG 6  
 DB 12 GGCTAG 17

```

RESULT 9
US-08-456-837-2
Sequence 2, Application US/08456837
Patent No. 5643774
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Gaffney, Phillip E.
APPLICANT: Hamm, Stephen Ting
APPLICANT: Hamm, Phillip E.
APPLICANT: Utnes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 05/08/456, 837
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8684
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-837-2

Query Match 90.9%; Score 30; DB 1; Length 538.
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GCGTSG 6
Db 12 GCGTAD 17

```

```

APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Lam, Stephen Ting
APPLICANT: Hamm, Phillip E.
APPLICANT: Utnes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8684
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-457-342-2

Query Match 90.9%; Score 30; DB 1; Length 538.
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GCGTSG 6
Db 12 GCGTNG 17

```

```

RESULT 10
US-08-457-342-2
Sequence 2, Application US/08457342
Patent No. 5639560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Hamm, Phillip E.
APPLICANT: Utnes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 05/08/456, 837
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8684
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-456-837-2

Query Match 90.9%; Score 30; DB 1; Length 538.
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GCGTSG 6
Db 12 GCGTAD 17

```

```

RESULT 11
US-08-457-646a-2
Sequence 2, Application US/08457646a
Patent No. 5639560
GENERAL INFORMATION:
APPLICANT: Schupp, Thomas
APPLICANT: Ligon, James M.
APPLICANT: Beck, James Joseph
APPLICANT: Hill, Dwight Steven
APPLICANT: Ryals, John Andrew
APPLICANT: Gaffney, Thomas Deane
APPLICANT: Gaffney, Phillip E.
APPLICANT: Hamm, Stephen Ting
APPLICANT: Hamm, Phillip E.
APPLICANT: Utnes, Scott Joseph
TITLE OF INVENTION: Genes for the synthesis of
antipathogenic substances
NUMBER OF SEQUENCES: 22
CORRESPONDENCE ADDRESS:
ADDRESSEE: Ciba-Geigy Corporation
STREET: 7 Skyline Drive
CITY: Hawthorne
STATE: NY
COUNTRY: USA
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: 01-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/457,205
FILING DATE: 01-JUN-1995
APPLICATION NUMBER: 08/258,261
FILING DATE: 08-JUN-1994
ATTORNEY/AGENT INFORMATION:
NAME: Elmer, James Scott
REGISTRATION NUMBER: 36,129
REFERENCE/DOCKET INFORMATION:
TELECOMMUNICATION INFORMATION:
TELEPHONE: 919-541-8614
TELEFAX: 919-541-8684
INFORMATION FOR SEQ ID NO.: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-457-342-2

Query Match 90.9%; Score 30; DB 1; Length 538.
Best Local Similarity 83.3%; Pred. No. 7.9e+02;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Cy 1 GCGTSG 6
Db 12 GCGTNG 17

```

STREET: 7 Skyline Drive  
 CITY: Hawthorne  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 FILING DATE: 01-JUN-1995  
 CLASSIFICATION: 530  
 PRIORITY NUMBER: 08/457,646A  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/457,205  
 FILING DATE: 01-JUN-1995  
 APPLICATION NUMBER: 08/258,261  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 538 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-457-646A-2

Query Match 90.9% Score 30: DB 1: Length 538;  
 Best Local Similarity 83.3% Pred. No. 7.9e+02;  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GGCTSG 6  
 DB 12 GGCTAG 17

RESULT 12  
 US-08-458-076A-2  
 Sequence 2, Application US/08458076A  
 Patent No. 56984250A  
 GENERAL INFORMATION:  
 APPLICANT: Schupp, Thomas  
 APPLICANT: Ligon, James M.  
 APPLICANT: Beck, James Joseph  
 APPLICANT: Hill, Dwight Steven  
 APPLICANT: Ryals, John Andrew  
 APPLICANT: Gaffney, Thomas Deane  
 APPLICANT: Lam, Stephen Ting  
 APPLICANT: Hamner, Phillip E.  
 APPLICANT: Uknes, Scott Joseph  
 TITLE OF INVENTION: Genes for the synthesis of  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ciba-Geigy Corporation  
 STREET: Skyline Drive  
 CITY: Hawthorne  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/458,076A  
 FILING DATE: 01-JUN-1995  
 CLASSIFICATION: 530  
 PRIORITY NUMBER: 08/457,205  
 APPLICATION NUMBER: 08/258,261  
 FILING DATE: 01-JUN-1995  
 APPLICATION NUMBER: 08/258,261  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
 TELEPHONE: 919-541-8614  
 TELEFAX: 919-541-8689  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 538 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-458-076A-2

Query Match 90.9% Score 30: DB 1: Length 538;  
 Best Local Similarity 83.3% Pred. No. 7.9e+02;  
 Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GGCTSG 6  
 DB 12 GGCTAG 17

RESULT 13  
 US-08-457-335A-2  
 Sequence 2, Application US/08457335A  
 Patent No. 5723759  
 GENERAL INFORMATION:  
 APPLICANT: Schupp, Thomas  
 APPLICANT: Ligon, James M.  
 APPLICANT: Beck, James Joseph  
 APPLICANT: Hill, Dwight Steven  
 APPLICANT: Ryals, John Andrew  
 APPLICANT: Gaffney, Thomas Deane  
 APPLICANT: Lam, Stephen Ting  
 APPLICANT: Hamner, Phillip E.  
 TITLE OF INVENTION: Genes for the synthesis of  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Ciba-Geigy Corporation  
 STREET: Skyline Drive  
 CITY: Hawthorne  
 STATE: NY  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 FILING DATE: 01-JUN-1995  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/457,205  
 FILING DATE: 01-JUN-1995  
 APPLICATION NUMBER: 08/258,261  
 ATTORNEY/AGENT INFORMATION:

NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8614  
FAX: 919-541-8615  
INFORMATION FOR SEQ ID NO:  
SEQUENCE CHARACTERISTICS: 2:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-457-35A-2

Query Match 90.9% Score 30: DB 1: Length 538:  
Best Local Similarity 83.3%: Pred. No. 7.9e+02:  
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
QY 1 GCGTSG 6  
DB 12 GCGTSG 17

RESULT 14  
US-08-729-214-2  
Sequence 2, Application US/08729214  
Patent No. 5817502  
GENERAL INFORMATION:  
APPLICANT: Ligou, James M.  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Hammer, Phillip E.  
APPLICANT: van Pee, Karl-Heinz  
APPLICANT: Kirner, Sabine  
TITLE OF INVENTION: genes for the synthesis of  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Ciba-Geigy Corporation  
STREET: 520 White Plains Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/729,214  
FILING DATE: TBA  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Young, Thomas R.  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
FAX: 919-541-8689  
INFORMATION FOR SEQ ID NO:  
SEQUENCE CHARACTERISTICS: 2:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-729-214-2

Query Match 90.9% Score 30: DB 2: Length 538:  
Best Local Similarity 83.3%: Pred. No. 7.9e+02:  
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:

QY 1 GCGTSG 6  
DB 12 GCGTSG 17

RESULT 15  
US-08-729-214-24  
Sequence 24 Application US/08729214  
Patent No. 5817502  
GENERAL INFORMATION:  
APPLICANT: Ligou, James M.  
APPLICANT: Hill, Dwight Steven  
APPLICANT: Ryals, John Andrew  
APPLICANT: Hammer, Phillip E.  
APPLICANT: van Pee, Karl-Heinz  
APPLICANT: Kirner, Sabine  
TITLE OF INVENTION: genes for the synthesis of  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESSES:  
ADDRESS: Ciba-Geigy Corporation  
STREET: 520 White Plains Road  
CITY: Tarrytown  
STATE: NY  
COUNTRY: USA  
ZIP: 10591  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/729,214  
FILING DATE: TBA  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Young, Thomas R.  
REGISTRATION NUMBER: 38,241  
REFERENCE/DOCKET NUMBER: CGC 1506/CIP5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-8587  
FAX: 919-541-8689  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-729-214-24

Query Match 90.9% Score 30: DB 2: Length 538:  
Best Local Similarity 83.3%: Pred. No. 7.9e+02:  
Matches 5: Conservative 1: Mismatches 0: Indels 0: Gaps 0:  
QY 1 GCGTSG 6  
DB 12 GCGTSG 17

RESULT 16  
US-09-028-934-2  
Sequence 2, Application US/09028934  
Patent No. 611870  
GENERAL INFORMATION:  
APPLICANT: Ligou, James M.  
APPLICANT: Hill, Dwight S.  
APPLICANT: Ryals, John Andrew  
APPLICANT: Hammer, Phillip E.  
APPLICANT: van Pee, Karl-Heinz  
APPLICANT: Kirner, Sabine  
APPLICANT: Young, Thomas R.

TITLE OF INVENTION: Pyroglutinin Biosynthesis Genes and Uses  
TITLE OF INVENTION: Pyroglutinin Biosynthesis Genes and Uses  
NUMBER OF SEQUENCES: 37  
CORRESPONDENCE ADDRESS:  
ADDRESS: No. 6117670artis Corporation  
STREET: 3054 Cornwallis Road  
CITY: Research Triangle Park  
STATE: NC  
COUNTRY: USA  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US 08/729, 214  
FILING DATE: 09-OCT-1996  
PRIORITY APPLICATION NUMBER: US 08/258, 261  
ATTORNEY/AGENT INFORMATION:  
NAME: Melgs, J. Timothy  
REGISTRATION NUMBER: 38, 241  
REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-5687  
FAX: 919-541-5689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-028-934-2

Query Match 90.9% Score 30; DB 3; Length 538;  
Best Local Similarity 83.3% Pred No. 7.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTSG 6  
||||:|  
DB 12 GCGTAC 17

RESULT 17  
US-09-642-660-10.ra1  
Sequence 24, Application US/09028934  
Patent No. 6117670  
GENERAL INFORMATION:  
APPLICANT: Ligou, James M.  
APPLICANT: Hill, Dwight S.  
APPLICANT: Lam, Steven T.  
APPLICANT: James, Philip D. E.  
APPLICANT: van Pelt, Karl-Heinz  
APPLICANT: Kimer, Sabine  
APPLICANT: Young, Thomas R.  
TITLE OF INVENTION: Pyroglutinin Biosynthesis Genes and Uses  
TITLE OF INVENTION: Pyroglutinin Biosynthesis Genes and Uses  
NUMBER OF SEQUENCES: 37  
CURRENT APPLICATION DATA: US 08/729, 214  
FILING DATE: 09-OCT-1996  
PRIORITY APPLICATION NUMBER: US 08/258, 261  
ATTORNEY/AGENT INFORMATION:  
NAME: Melgs, J. Timothy  
REGISTRATION NUMBER: 38, 241  
REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-5687  
FAX: 919-541-5689  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-028-934-2

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA: US 08/729, 214  
FILING DATE: 09-OCT-1996  
PRIORITY APPLICATION NUMBER: US 08/258, 261  
ATTORNEY/AGENT INFORMATION:  
NAME: Melgs, J. Timothy  
REGISTRATION NUMBER: 38, 241  
REFERENCE/DOCKET NUMBER: CGC1506/CIP7  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 919-541-5687  
FAX: 919-541-5689  
INFORMATION FOR SEQ ID NO: 24:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 538 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-028-934-24

Query Match 90.9% Score 30; DB 3; Length 538;  
Best Local Similarity 83.3% Pred No. 7.9e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTSG 6  
||||:|  
DB 12 GCGTAC 17

RESULT 18  
US-09-199-229-2  
Sequence 2, Application US/09199229  
Patent No. 603307  
GENERAL INFORMATION:  
APPLICANT: Yaver, Debbie S.  
APPLICANT: Berk, Randy M.  
APPLICANT: Rey, Michael W.  
TITLE OF INVENTION: Polypeptides Having Choline Oxidase  
TITLE OF INVENTION: Polypeptides Having Choline Oxidase  
TITLE OF INVENTION: Polypeptides Having Choline Oxidase  
TITLE OF INVENTION: Polypeptides Having Choline Oxidase  
CURRENT FILING DATE: 1998-11-24  
NUMBER OF SEQ ID NOS: 2  
SOFTWARE: FASTSO for Windows Version 3.0  
SEQ ID NO: 2  
LENGTH: 543  
TOPOLOGY: linear  
ORGANISM: Pseudomonas  
US-09-199-229-2

Query Match 90.9% Score 30; DB 3; Length 543;  
Best Local Similarity 83.3% Pred No. 8e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCGTSG 6  
||||:|  
DB 20 GCGTAC 25

RESULT 19  
US-09-443-087-2  
Sequence 2, Application US/09443087



Patent No. 6146864  
 GENERAL INFORMATION:  
 APPLICANT: Debole Yaver  
 APPLICANT: Debole Yaver  
 APPLICANT: Michael A. Berka  
 TITLE OF INVENTION: Polypeptides Having Choline Oxidase  
 TITLE OF INVENTION: Activity And Nucleic Acids Encoding Same  
 FILE REFERENCE: 5735,200-US  
 CURRENT APPLICATION NUMBER: US/09/443,087  
 CURRENT FILING DATE: 1999-11-16  
 EARLIER FILING DATE: 1999-11-16  
 EARLIER FILING DATE: 1999-11-24  
 NUMBER OF SEQ ID NOS: 2  
 SOFTWARE: Pstseq for Windows Version 3.0  
 SEQ ID NO 2  
 LENGTH: 543  
 TYPE: PRT  
 FUNCTION: Pusarium venenatum  
 US-09-443-087-2

Query Match 90.9%; Score 30; DB 4; Length 543;  
 Best Local Similarity 83.3%; Pred. No. 8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 20 GCGTGG 25

RESULT 20  
 US-09-648A-45  
 Sequence 445 Application US/08955648A  
 Patent No. 5565105  
 GENERAL INFORMATION:  
 APPLICANT: Mc Wherter, Charles  
 APPLICANT: Peng, Yiqing  
 TITLE OF INVENTION: No. 5969105e1 Stem Cell Factor Receptor  
 TITLE OF INVENTION: Agonists  
 NUMBER OF SEQUENCES: 6  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: G. D. Seattle & Co.  
 STREET: P.O. Box 5110  
 CITY: Chicago  
 STATE: IL  
 COUNTRY: U. S. A.  
 ZIP CODE: 60680  
 COMPUTER RELEASABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 PRIORITY ACTION NUMBER: US/08/955,848A  
 FILING DATE: 1997-11-22  
 CLASSIFICATION: 516  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 60/029,165  
 FILING DATE: 25-OCT-1997  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Benmuel, Dennis A.  
 REGISTRATION NUMBER: 33,924  
 REFERENCE/DOCKET INFORMATION:  
 TELEPHONE: 314-737-6986  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 314-737-6986  
 TELEFAX: 314-737-6972  
 ELEX:  
 IMPROVED FOR SEQ ID NO: 45;  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: Single  
 TOPOLOGY: linear

US-08-955-848A-45

Query Match 87.9%; Score 29; DB 2; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 47;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 5 GCGTGG 10

RESULT 21  
 US-08-501-253A-9  
 Sequence 9 Application US/08501253A  
 Patent No. 6146638  
 GENERAL INFORMATION:  
 APPLICANT: Uckun, Fatih  
 APPLICANT: Tumer, Nilgun  
 TITLE OF INVENTION: Biotherapeutic Agents Comprising  
 TITLE OF INVENTION: Recombinant PAP and PAP Mutants  
 NUMBER OF SEQUENCES: 21  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Merchante & Gould  
 STREET: 90 South 7th Street, 3100 No. 6146628west Center  
 CITY: Minneapolis  
 STATE: MN  
 COUNTRY: US  
 ZIP CODE: 55402  
 COMPUTER RELEASABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 PRIORITY ACTION NUMBER: US/08/501,253A  
 FILING DATE: 1995-11-09  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kettleberger, Denise M.  
 REGISTRATION NUMBER: 33,924  
 REFERENCE/DOCKET INFORMATION:  
 TELEPHONE: 612-312-9081  
 TELEFAX: 612-312-9081  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 15 amino acids  
 TYPE: amino acid  
 MEDIUM: linear  
 TOPOLOGY: linear  
 US-08-501-253A-9

Query Match 87.9%; Score 29; DB 4; Length 15;  
 Best Local Similarity 83.3%; Pred. No. 47;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 GCGTSG 6  
 |||||  
 Db 7 GCGTGG 12

RESULT 22  
 US-08-256-156A-2  
 Sequence 9 Application US/08256156A  
 Patent No. 5817821  
 GENERAL INFORMATION:  
 APPLICANT: Wu, Anna M.  
 APPLICANT: Wu, Anna M.  
 TITLE OF INVENTION: No. 5837821e1 Antibody Constructs  
 NUMBER OF SEQUENCES: 2  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Notinwell, Flagg, Ernst & Kurz, P.C.

STREET: Suite 701 East Tower, 555 Thirteenth  
 STREET: Street, N.W.  
 CITY: Washington, D.C.  
 STATE: none  
 COUNTRY: United States  
 ZIP: 20004  
 COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/256,156A  
 FILING DATE: 24-JAN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Ernst, Barbara G.  
 REGISTRATION NUMBER: 30,377  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-383-6040  
 TELEFAX: 202-383-6040  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 24 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 MOLECULE TYPE: peptide  
 US-08-256-156A.2

Query Match 87.9% Score 29; DB 2; Length 24;  
 Best Local Similarity 83.3% Pred. No. 71;

Matches 5; Conservative 1; Mismatches 0; Gaps 0;

1 GCGTSG 6  
 |||||  
 Db 14 GCGSSG 19

RESULT 23  
 US-08-377-687-39  
 Patent No. 5538525  
 GENERAL INFORMATION:  
 APPLICANT: BROEKERT, WILLEM F.  
 APPLICANT: CAMME, BRUNO P.A.  
 APPLICANT: OSMORN, RUPERT W.  
 APPLICANT: RESS, SMARH B. G.  
 APPLICANT: THERAS, FRANK R.G.  
 APPLICANT: VANDERLEYDEN, JOZEF  
 TITLE OF INVENTION: BIOCIDAL PROTEINS  
 NUMBER OF SEQUENCES: 59  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DABRY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON  
 STATE: DC  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/377,687  
 FILING DATE:  
 CLASSIFICATION: 800  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/002,480  
 FILING DATE: 04-JAN-1993

ATTORNEY/AGENT INFORMATION:  
 NAME: KOKULIS, PAUL N.  
 REGISTRATION NUMBER: 16,773  
 REFERENCE/DOCKET NUMBER: 99042/SER. 36525/US/A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-822-0940  
 TELEFAX: 202-822-0944  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 42 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 MOLECULE TYPE: linear  
 US-08-377-687-39

Query Match 87.9% Score 29; DB 1; Length 42;  
 Best Local Similarity 83.3% Pred. No. 1,2e+02;  
 Matches 5; Conservative 1; Mismatches 0; Gaps 0;

1 GCGTSG 6  
 |||||  
 Db 24 GCGSSG 29

RESULT 24  
 US-08-377-192-39  
 Patent No. 5824869  
 GENERAL INFORMATION:  
 APPLICANT: BROEKERT, WILLEM F.  
 APPLICANT: CAMME, BRUNO P.A.  
 APPLICANT: OSMORN, RUPERT W.  
 APPLICANT: THERAS, FRANK R.G.  
 APPLICANT: VANDERLEYDEN, JOZEF  
 TITLE OF INVENTION: BIOCIDAL PROTEINS  
 NUMBER OF SEQUENCES: 59  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: CUSHMAN DABRY & CUSHMAN  
 STREET: 1100 NEW YORK AVENUE, N.W.  
 CITY: WASHINGTON  
 STATE: D.C.  
 COUNTRY: USA  
 ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/777,192  
 FILING DATE:  
 CLASSIFICATION:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/002,480  
 FILING DATE: 04-JAN-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KOKULIS, PAUL N.  
 REGISTRATION NUMBER: 16,773  
 REFERENCE/DOCKET NUMBER: 99042/SER. 36525/US/A  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-822-0940  
 TELEFAX: 202-822-0944  
 INFORMATION FOR SEQ ID NO: 39:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 42 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 MOLECULE TYPE: linear  
 US-08-377-192-39

Query Match 87.9%: Score 29; DB 2; Length 42;  
Best Local Similarity 83.3%: Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTSG 6  
111111  
24 GGCTSG 29

RESULT 25  
US-08-971-982-39  
Sequence 39, Application US/08971982  
Patent No. 6187904  
GENERAL INFORMATION:  
APPLICANT: HROENKERT, MILLEN F.  
NAME: BRUNO P.A.  
ADDRESS: 10000 10TH AVE.  
CITY: WASHINGTON  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005  
COMPUTER READABLE FORM:  
FILE TYPE: Floppy disk  
COMPUTER SYSTEM: PC-DOS/MS-DOS  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/971,982  
FILING DATE: 17-NOV-1997  
CLASSIFICATION: C01K 01/00  
PRIORITY APPLICATION: 04-JAN-1999  
APPLICATION NUMBER: US 08/002,480  
FILING DATE: 04-JAN-1999  
ATTORNEY/AGENT INFORMATION:  
NAME: KOKULIS, PAUL N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: 95042/SEE.16525/US/A  
TELEPHONE: 202-165-1200  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
SEQUENCE DESCRIPTION: SEQ ID NO: 39:  
US-08-971-982-39

Query Match 87.9%: Score 29; DB 4; Length 42;  
Best Local Similarity 83.3%: Pred. No. 1.2e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTSG 6  
111111  
24 GGCTSG 29

RESULT 26  
PCT-US96-01720-10  
Sequence 10, Application PC/TUS9601720

GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01720  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
NAME: Parmelee, Steven W.  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-5PC  
INFORMATION FOR SEQ ID NO: 10:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-01720-10

Query Match 87.9%: Score 29; DB 5; Length 71;  
Best Local Similarity 83.3%: Pred. No. 1.9e+02;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGCTSG 6  
111111  
59 GGCTSG 64

RESULT 27  
PCT-US96-01720-11  
Sequence 11, Application PC/TUS9601720  
GENERAL INFORMATION:  
TITLE OF INVENTION: MODIFIED-AFFINITY STREPTAVIDIN  
NUMBER OF SEQUENCES: 11  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US96/01720  
FILING DATE:  
CLASSIFICATION:  
PRIORITY APPLICATION DATA:  
NAME: Parmelee, Steven W.  
FILING DATE: 09-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Parmelee, Steven W.  
REGISTRATION NUMBER: 31,990  
REFERENCE/DOCKET NUMBER: 16336-5PC  
INFORMATION FOR SEQ ID NO: 11:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 42 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US96-01720-11

Query Match 87.9%: Score 29; DB 5; Length 71;

Best Local Similarity: 83.3%; Pred. No. 1.9e+02;  
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTSG 6  
DB 59 GGGTGG 64

## RESULT 28

US-08-325-253-10  
Sequence 10, Application US/08325253  
Patent No. 5686564  
OPERATING SYSTEM: IBM PC compatible  
APPLICANT: Derek Edward Brundish  
APPLICANT: Hans Rink  
APPLICANT: Markus Griffler  
APPLICANT: John Peter Pfleistle  
TITLE OF INVENTION: Peptide Derivatives Corresponding to  
Patent No. 5686564  
DESCRIPTION: The Carboxy Terminal Sequence of Mitridin  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: P.O. Box 12257  
CITY: Research Triangle Park  
STATE: NC 27709  
COUNTRY: US  
ZIP: 27709  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy Disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
APPLICATION NUMBER: US/08/325,253  
FILING DATE: 20-OCT-1994  
CLASSIFICATION: 510  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/EP 93/00908  
FILING DATE: 13-Apr-11-1993  
PRIORITY INFORMATION: England 9209032.3  
APPLICATION NUMBER: 25-Apr-11-1992  
FILING DATE: 25-Apr-11-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: W. Murray Spruill  
REGISTRATION NUMBER: 32,943  
REFERENCE/DOCKET NUMBER: 4-18956/A  
INVENTOR: W. Murray Spruill  
SEQUENCE CHARACTERISTICS:  
LENGTH: 91 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-325-253-10

Query Match 87.9%; Score 29; DB 1; Length 91;  
Best Local Similarity: 83.3%; Pred. No. 2.3e+02;  
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GGGTSG 6  
DB 71 GGGTGG 76

## RESULT 29

US-08-481-6588-50  
Sequence 50, Application US/084816588  
Patent No. 5955075  
GENERAL INFORMATION:  
APPLICANT: Zavarok, Jan  
APPLICANT: Pastorekova, Silvia

APPLICANT: Pastorek, Jaronlt  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920

COMPUTER READABLE FORM:  
MEDIUM TYPE: IBM PC compatible  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/481,6588  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: 0-0-0021.3E  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: Region of homology to collagen alpha  
US-08-481-6588-50

Query Match 87.9%; Score 29; DB 2; Length 98;  
Best Local Similarity: 83.3%; Pred. No. 2.5e+02;  
Matches 5: Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGTSG 6  
DB 13 GGGSTG 18

RESULT 30  
US-08-477-504A-50  
Sequence 50, Application US/08477504A  
Patent No. 5972353  
GENERAL INFORMATION:  
APPLICANT: Zavarok, Jan  
APPLICANT: Pastorek, Jaronlt  
APPLICANT: Pastorek, Jaronlt  
TITLE OF INVENTION: MN Gene and Protein  
NUMBER OF SEQUENCES: 86  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
APPLICATION NUMBER: US/08/477,504A

FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 1.0  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/260,190  
 FILING DATE: 15-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lauder, Leona L.  
 REGISTRATION NUMBER: 30,863  
 REFERENCE/DOCKET NUMBER: D-0021.3D  
 TELEPHONE: 415-435-2034  
 TELEFAX: 415-435-0727  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 98 amino acids  
 TYPE: CDS  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 DESCRIPTION: Region of homology to collagen alpha  
 US-08-47-504A-50

Query Match 87.9% Score 29; DB 2; Length 98;  
 Best Local Similarity 83.3% Pred. No. 2.5e+02:  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 111111  
 Db 13 GGGSSG 18

RESULT 31  
 US-08-486-756A-50  
 Sequence 50, Application US/08486756A  
 Patent No. 5981711  
 GENERAL INFORMATION:  
 APPLICANT: Zavotekova, Silvia  
 APPLICANT: Pastorek, Jaromir  
 TITLE OF INVENTION: MN Gene and Protein  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Leona L. Lauder  
 STREET: 6 Mariposa Court  
 CITY: Tiburon  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94920  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30 (ERO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/486,756A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/260,190  
 FILING DATE: 15-JUN-1994  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lauder, Leona L.  
 REGISTRATION NUMBER: 30,863  
 REFERENCE/DOCKET NUMBER: D-0021.3C  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-435-2034  
 TELEFAX: 415-435-0727  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 98 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 DESCRIPTION: Region of homology to collagen alpha  
 US-08-486-756A-50

Query Match 87.9% Score 29; DB 2; Length 98;  
 Best Local Similarity 83.3% Pred. No. 2.5e+02:  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 111111  
 Db 13 GGGSSG 18

RESULT 32  
 US-08-485-862B-50  
 Sequence 50, Application US/08485862B  
 Patent No. 5989838  
 GENERAL INFORMATION:  
 APPLICANT: Zavotekova, Silvia  
 APPLICANT: Pastorek, Jaromir  
 TITLE OF INVENTION: MN Gene and Protein  
 NUMBER OF SEQUENCES: 86  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Leona L. Lauder  
 STREET: 6 Mariposa Court  
 CITY: Tiburon  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94920  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30 (ERO)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/485,862B  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/477,504  
 FILING DATE: 07-JUN-1995  
 APPLICATION NUMBER: 15-JUN-1994  
 FILING DATE: 15-JUN-1994  
 APPLICATION NUMBER: US 08/260,190  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Lauder, Leona L.  
 REGISTRATION NUMBER: 30,863  
 REFERENCE/DOCKET NUMBER: D-0021.3D  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 415-435-2034  
 TELEFAX: 415-435-0727  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 98 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 DESCRIPTION: Region of homology to collagen alpha  
 US-08-485-862B-50

Query Match 87.9% Score 29; DB 2; Length 98;  
 Best Local Similarity 83.3% Pred. No. 2.5e+02:  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 Oy 1 GGGTSG 6  
 111111  
 Db 13 GGGSSG 18

RESULT 33  
US-08-487-077A-50  
Sequence 50: Application US/08487077A  
Patent No. 6069342  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorek, Jatomir  
TITLE OF INVENTION: MN Gene and Protein  
CLASSIFICATION: 3514  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487.077A  
FILING DATE: 07-JUN-1995  
PRIORITY INFORMATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
TELEPHONE: 415-435-2034  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: Region of homology to collagen alpha  
DESCRIPTION: 1 chain  
US-08-487-077A-50

Query Match 87.9%; Score 29; DB 3; Length 98;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTSG 6  
DB 13 GCGSSG 18  
|||||

RESULT 34  
US-08-485-863A-50  
Sequence 50: Application US/08485863A  
Patent No. 6093548  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
TITLE OF INVENTION: MN Gene and Protein  
CLASSIFICATION: 3514  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 6 Mariposa Court  
CITY: Tiburon  
STATE: California  
COUNTRY: USA  
ZIP: 94920

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485.863A  
FILING DATE: 07-JUN-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3G  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-435-2034  
TELEFAX: 415-435-0727  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 98 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
DESCRIPTION: Region of homology to collagen alpha  
DESCRIPTION: 1 chain  
US-08-485-863A-50

Query Match 87.9%; Score 29; DB 3; Length 98;  
Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 CGCTSG 6  
DB 13 GCGSSG 18  
|||||

RESULT 35  
US-08-485-049D-50  
Sequence 50: Application US/08485049D  
Patent No. 6204370  
GENERAL INFORMATION:  
APPLICANT: Zavada, Jan  
APPLICANT: Pastorekova, Silvia  
TITLE OF INVENTION: MN Gene and Protein  
CLASSIFICATION: 3514  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Leona L. Lauder  
STREET: 369 Pine Street  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94104  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485.049D  
FILING DATE: 07-JUN-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/260,190  
FILING DATE: 15-JUN-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Lauder, Leona L.  
REGISTRATION NUMBER: 30,863  
REFERENCE/DOCKET NUMBER: D-0021.3E  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-981-2034  
 TELEFAX: 415-981-0332  
 INFORMATION FOR SEQ ID NO: 50:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 98 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 DESCRIPTION: Region of homology to collagen alpha  
 US-08-485-043D-58

Query Match 87.9%; Score 29; DB 4; Length 98;  
 Best Local Similarity 83.3%; Pred. No. 2.5e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Or 1 GCGTSG 6  
 |||||  
 Db 13 GGGSSG 18

RESULT 36  
 US-08-905-223-319  
 SEQ ID NO: 319 Application US/08905223  
 Patent No. 6220029

GENERAL INFORMATION:

APPLICANT: Edwards, Jean-Baptiste D.

APPLICANT: Dugier, Aymeric

APPLICANT: Lacroix, Bruno

TITLE OF INVENTION: 5' ESTS FOR SECRETED PROTEINS

NUMBER OF SEQUENCES: 503

CORRESPONDENCE ADDRESS

ADDRESSER: Knobbe, Martens, Olson & Bear

STREET: 501 West Broadway

CITY: San Diego

STATE: California

COUNTRY: USA

FILE NO: 92101-3505

COMPUTER: IBM PC COMPATIBLE

MEDIUM TYPE: Floppy Disk

OPERATING SYSTEM: WIN95

SOFTWARE: Word

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/905,223

CLASSIFICATION: 536

ATTORNEY/AGENT INFORMATION:

NAME: Israel, Nod A.

REGISTRATION NUMBER: 29,655

REFERENCE/DOCKET NUMBER:

TELEPHONE: (619) 235-8550

TELEFAX: (619) 235-8550

INFORMATION FOR SEQ ID NO: 319:

SEQUENCE CHARACTERISTICS:

LENGTH: 105 amino acids

TYPE: AMINO ACID

TOPOLOGY: LINEAR

MOLCULE TYPE: PROTEIN

ORGANISM: Homo Sapiens

TISSUE TYPE: Brain

FEATURE:

NAME/KEY: sig-peptide

LOCATION: -25..-1

IDENTIFICATION METHOD: Von Heljje matrix

IDENTIFICATION: score 5.7

OTHER INFORMATION: seq STRFLAVAGPMA/RP

Query Match 87.9%; Score 29; DB 6; Length 105;  
 Best Local Similarity 83.3%; Pred. No. 2.6e+02;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Or 1 GCGTSG 6  
 |||||  
 Db 69 GGGTSG 74

RESULT 37  
 US-07-596-7008-107  
 SEQ ID NO: 107 Application US/075967008  
 Patent No. 5252466

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

APPLICANT: CRONN, JOHN E.

STRANDEDNESS: Single  
TOPOLOGY: Linear  
MOLECULE TYPE: Peptide  
US-07-956-700h-107

Query Match  
Best Local Similarity 83.3%; Pred. No. 3e+02;  
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 32 GCGTGC 37

RESULT 39  
US-08-476-537-107  
Sequence 107, Application US/08476537  
PATENT No. 5756290  
GENERAL INFORMATION:  
APPLICANT: Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCE: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durfee  
STREET: 321 No. 5756290th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08476,537  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756290thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-744-0090  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 123 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
MOLECULE TYPE: Peptide  
TOPOLOGY: Linear  
US-08-476-537-107

Query Match  
Best Local Similarity 87.9%; Score 29; DB 1; Length 123;  
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 32 GCGTGC 37

RESULT 40  
US-08-485-607-107  
Sequence 107, Application US/08485607  
PATENT No. 5756297

GENERAL INFORMATION:  
APPLICANT: Robert Haselkorn and Piotr Gornicki  
TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
TITLE OF INVENTION: Carboxylase  
NUMBER OF SEQUENCE: 116  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Arnold, White & Durfee  
STREET: 321 No. 5756297th Clark Street  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60610  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC Compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: ASCII-DOS  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/485,607  
FILING DATE: 07/956-195  
CLASSIFICATION: 400  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/956,700  
FILING DATE: 10/21/92  
ATTORNEY/AGENT INFORMATION:  
NAME: Thomas E. No. 5756271thrup  
REGISTRATION NUMBER: 33,268  
REFERENCE/DOCKET NUMBER: ARCD:058  
TELEPHONE: 1-312-744-0090  
TELEFAX: 1-312-755-4489  
INFORMATION FOR SEQ ID NO: 107:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 123 amino acids  
TYPE: Amino acid  
STRANDEDNESS: Single  
MOLECULE TYPE: Peptide  
US-08-485-607-107

Query Match  
Best Local Similarity 87.9%; Score 29; DB 1; Length 123;  
Matches 5: Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 32 GCGTGC 37

Search completed: February 4, 2002, 08:01:36  
Job time: 80 sec



---

**THIS PAGE BLANK (USPTO)**

Gencore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2002, 08:00:16 ; Search time 23.49 Seconds

(without alignments)  
19.457 Million cell updates/sec

Title: US-09-642-660-10

Perfect score: 3.56

Sequence: 1 GCGTGG 6

Scoring table: BLASTM62

Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 7617453 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 100 summaries

Database : 1: p1r1.\*

2: p1r2.\*

3: p1r3.\*

4: p1r4.\*

Prod. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	100.0	102	2 S00729	hypothetical prote
2	33	100.0	234	2 IS4074	upstream stimulat
3	33	100.0	279	2 T26555	hypothetical prote
4	33	100.0	289	2 T01227	probable GTP-like
5	33	100.0	303	2 T48493	hypothetical prote
6	33	100.0	304	2 A57215	glial cells missin
7	33	100.0	304	2 A57215	glial cells missin
8	33	100.0	358	2 B23456	phosphotransfere
9	33	100.0	358	2 B23456	phosphotransfere
10	33	100.0	569	2 S32494	glucokinase regula
11	33	100.0	572	2 B28474	phosphotransfere
12	33	100.0	574	2 T50766	mandelonitrile ly
13	33	100.0	578	2 B23697	phosphotransfere
14	33	100.0	578	2 T48795	origin recognition
15	33	100.0	625	2 E57482	hypothetical prote
16	33	100.0	627	2 S41745	glucokinase regula
17	33	100.0	629	2 A29661	Keratin, 68k type
18	33	100.0	658	1 Q08E61	U52 protein - hum
19	33	100.0	760	2 T27941	hypothetical prote
20	33	100.0	923	1 B35905	endopeptidase CIP
21	33	100.0	1125	1 T1251	endopeptidase CIP
22	33	100.0	1130	2 B23669	30S ribosomal prot
23	30	90.9	136	1 S70890	hypothetical prote
24	30	90.9	136	2 D65915	hypothetical prote
25	30	90.9	143	2 T46122	30S ribosomal prot
26	30	90.9	185	2 T46122	hypothetical prote
27	30	90.9	185	2 T46122	hypothetical prote
28	30	90.9	245	2 S55827	conserved hypote
29	30	90.9	255	2 T03371	glycine-rich prote

30	90.9	276	2 F75300	hypothetical prote
31	90.9	278	2 G85887	hypothetical prote
32	90.9	278	2 B65070	ethanimine UTP
33	90.9	278	2 B65070	ethanimine UTP
34	90.9	340	2 T20807	hypothetical prote
35	90.9	364	2 T34954	probable UDP-N-ac
36	90.9	370	2 I57555	c-Maf protein - mo
37	90.9	372	2 B34261	alanine dehydrogen
38	90.9	382	1 B46233	transcription fact
39	90.9	382	2 T48109	hypothetical prote
40	90.9	382	2 T48109	hypothetical prote
41	90.9	385	2 S42680	phosphoserine turn
42	90.9	396	2 A12126	D-alanine-D-alani
43	90.9	410	2 E70579	probable nuc prot
44	90.9	453	2 B70893	hypothetical glycy
45	90.9	457	2 C81138	lipoamide dehydro
46	90.9	467	2 A81881	probable serine/th
47	90.9	480	2 A84427	probable serine/th
48	90.9	487	2 G64235	hypothetical prote
49	90.9	487	2 G64235	hypothetical prote
50	90.9	501	2 F86275	hypothetical prote
51	90.9	503	2 E86275	hypothetical prote
52	90.9	527	2 C70397	periplasmic cell d
53	90.9	544	2 T44938	transducer protein
54	90.9	552	2 A56756	hypothetical prote
55	90.9	557	2 T47723	mandelonitrile ly
56	90.9	572	2 C81359	probable oxidoredu
57	90.9	576	2 T08073	mandelonitrile ly
58	90.9	589	2 T50698	probable mandelon
59	90.9	594	2 T50765	adhesion of calyx
60	90.9	594	2 T50765	adhesion of calyx
61	90.9	667	2 A70883	hypothetical glycy
62	90.9	667	2 A70883	hypothetical glycy
63	90.9	806	2 T13630	hypothetical prote
64	90.9	806	2 T13630	hypothetical prote
65	90.9	896	2 T51891	dead ringer nuclea
66	90.9	901	2 T56093	exonuclease ABC c
67	90.9	916	2 H72372	collagen alpha 3(V
68	90.9	959	2 S32605	probable ATP-bind
69	90.9	998	2 T35745	RNA-directed DNA p
70	90.9	1051	2 B27672	hypothetical prote
71	90.9	1106	2 A56673	hypothetical prote
72	90.9	1106	2 A56673	hypothetical prote
73	90.9	1148	2 S51655	hypothetical prote
74	90.9	1301	2 S81118	alpha-amyase - Al
75	90.9	2314	2 T28518	hypothetical prote
76	90.9	3429	2 T13857	hypothetical prote
77	90.9	3828	2 T13857	hypothetical prote
78	90.9	33	2 A22483	citronox protein
79	90.9	33	2 A22483	citronox protein
80	90.9	144	2 S33051	glycine rich prote
81	90.9	144	2 S33051	glycine rich prote
82	90.9	91	1 S00060	biotin carboxyl ca
83	90.9	125	2 T49366	M6067 homolog D0-
84	90.9	135	2 S73564	probable gas vasc
85	90.9	144	2 T34720	hypothetical prote
86	90.9	156	2 C86122	hypothetical prote
87	90.9	156	2 C86122	hypothetical prote
88	90.9	156	2 S72733	hypothetical prote
89	90.9	161	2 T46048	hypothetical prote
90	90.9	166	1 K80CB	Keratin, 68k type
91	90.9	167	2 S21359	hypothetical prote
92	90.9	177	2 F81442	hypothetical prote
93	90.9	183	2 F80109	Keratin, 68k type
94	90.9	183	2 F80109	Keratin, 68k type
95	90.9	201	1 A55012	signal peptidase 2
96	90.9	226	2 S26300	transcription fact
97	90.9	243	2 B33144	homeotic protein U
98	90.9	243	2 B33144	homeotic protein U
99	90.9	257	2 C84830	hypothetical prote
100	90.9	257	2 C84830	hypothetical prote

## ALIGNMENTS

RESULT 1

Hypothetical protein 2 - Methanococcus thermolithotrophicus (fragment)

C:Species: Methanococcus thermolithotrophicus

C:Date: 31-Dec-1988 #sequence\_revision 31-Dec-1988 #text\_change 22-Oct-1999

C:Accession: S00739

C:Title: Nucleotide sequence of regions homologous to nlfH (nitrogenase re protein) from

ery. Implications.

A:Accession: S00737; M01D:88239240

A:Residues: 1-102 &lt;SD&gt;

A:Note: the authors translated the codon CAG for residue 91 as His

Query Match

Best Local Similarity 100.0%; Score 33; DB 2; Length 102;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GCGTSC 6

DB 7 GCGTSC 12

RESULT 2

upstream stimulatory factor - human (fragment)

C:Species: Homo sapiens (mm)

C:Date: 02-Jul-1996 #sequence\_revision 02-Jul-1996 #text\_change 05-Nov-1999

C:Accession: I54074

C:Title: Members of the USF family of helix-loop-helix proteins bind DNA as homo-

A:Accession: I54074

A:Residues: 1-234 &lt;RES&gt;

A:Cross-references: GB:S50537; NID:g261070; PTDN:ABD24368.1; PID:g261071

Query Match

Best Local Similarity 100.0%; Score 33; DB 2; Length 234;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GCGTSC 6

DB 1 GCGTSC 6

RESULT 3

Hypothetical protein Y22F5A.5 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000

C:Accession: T26555

C:Title: Submitted to the EMBL Data Library, January 1998

A:Accession: T26555

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Residues: 1-279 &lt;LID&gt;

A:Cross-references: EMBL:AL021479; PTDN:CAA16324.1; GSPDB:GN00023; CESP:Y22F5A.5

A:Genetic: EMBL: clone T22F5A

A:Gene: CESP:Y22F5A.5

A:Map position: 5  
A:Introns: 142/3  
C:Superfamily: Caenorhabditis elegans hypothetical protein C02A12.4

Query Match

Best Local Similarity 100.0%; Score 33; DB 2; Length 279;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GCGTSC 6

DB 206 GCGTSC 211

RESULT 4

Probable GT-1-like transcription factor (imported) - Arabidopsis thaliana

N:Alternate names: hypothetical protein F16M14.18

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 12-Feb-1999 #sequence\_revision 12-Feb-1999 #text\_change 23-Mar-2001

C:Accession: U01257; M01D:58402 Lin, X.; Ketchum, K.A.; Crosby, M.L.; Brandon, R.C.; Syke

submitted to the EMBL Data Library, July 1998

A:Description: Arabidopsis thaliana chromosome II BAC F16M14 genomic sequence.

A:Reference number: Z14213

A:Accession: T01257

A:Status: translated from GB/EMBL/DBJ

A:Residues: 1-289 &lt;SD&gt;

A:Cross-references: EMBL:AC003028; NID:g3335356; PID:g3335373

A:Experimental sources: Cultivar Columbia

A:Residues: 1-289 &lt;SD&gt;

A:Cross-references: EMBL:AC003028; NID:g3335356; PID:g3335373

A:Experimental sources: Cultivar Columbia

A:Residues: 1-289 &lt;SD&gt;

A:Cross-references: GB:AB002093; NID:g3335373; PTDN:AMC27174.1; GSPDB:GN00139

A:Gene: F16M14.18; Atg38250

A:Map position: 2

A:Introns: 102/3

Query Match

Best Local Similarity 100.0%; Score 33; DB 2; Length 289;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 1 GCGTSC 6

DB 142 GCGTSC 147

RESULT 5

Hypothetical protein T28J14.160 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cross)

C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 15-Sep-2000

C:Accession: T48493

C:Title: Submitted to the Protein Sequence Database, April 2000

A:Accession: T48493

A:Status: preliminary

A:Residues: 1-303 &lt;REV&gt;

A:Cross-references: EMBL:AL163652

A:Experimental source: Cultivar Columbia; BAC clone T28J14

C:Genetics:

A:Map position: 5  
A:Accession: 67/1; 152/3; 201/3  
A:Title: 128014.100  
C:Superfamily: Atkidiopsis thaliana hypothetical protein P109039.6

Query Match  
Best Local Similarity 100.0%; Score 33; DB 2; Length 303;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GCGTSG 6  
Db 14 GCGTSG 19

RESULT 6  
A:7215 cellis missing (gcm) protein - fruit fly (Drosophila melanogaster)  
C:Species: Drosophila melanogaster  
C:Date: 10-Nov-1995 #sequence\_revision 10-Nov-1995 #text\_change 21-Jul-2000  
C:Accession: A57215; A57216  
R:Hosoya, T.; Takizawa, K.; Nitta, K.; Hotta, Y.  
Cell 82, 1025-1036, 1995  
A:Title: glial cells missing: a binary switch between neuronal and glial determination  
A:Reference number: A57215; MUID:96016097  
A:Status: not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-504 <HOS>  
A:Cross-references: GB:D64040; NID:91020400; PID:d1011557; PID:91020401  
R:Jones, B.W.; Felter, R.D.; Teat, G.; Goodman, C.S.  
Cell 82, 1013-1023, 1995  
A:Title: glial cells missing: a genetic switch that controls glial versus neuronal fate.  
A:Reference number: A57216; MUID:96016096  
A:Accession: A57215  
A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-504 <JON>  
A:Cross-references: GB:034039; NID:91041730; PID:NAC6912.1; PID:91041731  
A:Gene: gcm  
A:Cross-references: FlyBase:FBgn0041479  
C:Keywords: nucleus

Query Match  
Best Local Similarity 100.0%; Score 33; DB 2; Length 504;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GCGTSG 6  
Db 430 GCGTSG 435

RESULT 7  
A:32156  
C:Species: Prunus serotina (black cherry)  
C:Date: 22-Nov-1993 #sequence\_revision 13-Mar-1997 #text\_change 17-Mar-2000  
C:Accession: S32156  
R:Gibson, J.E.  
submitted to the Cambridge library  
A:Description: Nucleotide sequence of a full-length cDNA clone encoding Prunus serotina  
A:Reference number: S32156  
A:Accession: S32156  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Cross-references: CML:Y73617; NID:9288115; PID:9288116  
C:Superfamily: alcohol oxidase  
C:Keywords: aldehyde-lyase; carbon-carbon lyase

Query Match  
Best Local Similarity 100.0%; Score 33; DB 2; Length 563;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GCGTSG 6  
Db 60 GCGTSG 65

RESULT 8  
B23696  
C:Species: Lactococcus lactis  
C:Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 15-Oct-1999  
R:de Vos, W.M.; Boerlijet, I.; van Royen, R.J.; Relche, B.; Hengstenberg, W.  
J. Biol. Chem. 265, 22554-22560, 1990  
A:Title: Characterization of the lactose-specific enzymes of the phosphotransferase  
A:Reference number: A23696; MUID:91093107  
A:Accession: B23696  
A:Status: preliminary  
A:Molecule type: cDNA  
A:Residues: 1-563 <BO>  
A:Cross-references: GB:M60447; GB:057578; NID:9149404; PID:AAA25182.1; PID:9149411  
C:Keywords: phosphotransferase; transmembrane protein

Query Match  
Best Local Similarity 100.0%; Score 33; DB 2; Length 568;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GCGTSG 6  
Db 475 GCGTSG 480

RESULT 9  
S32494  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 25-Dec-1994 #sequence\_revision 26-May-1995 #text\_change 17-Mar-1999  
C:Accession: S32494  
R:Beethuis, J.; Vandekeerehove, J.; van Schaftingen, E.  
FEBS Lett. 331, 113-117, 1993  
A:Title: Cloning and sequencing of rat liver cDNAs encoding the regulatory protein of  
A:Reference number: S32494; MUID:93238935  
A:Accession: S32494  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-569 <DE>  
A:Cross-references: EMBL:X68497  
A:Notes: In the published translation residues 429-430 are shown after residue 450 and,  
C:Superfamily: glucokinase regulator

Query Match  
Best Local Similarity 100.0%; Score 33; DB 2; Length 569;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GCGTSG 6  
Db 106 GCGTSG 111

RESULT 10  
B28474  
C:Species: Staphylococcus aureus  
C:Date: 20-Jun-1989 #sequence\_revision 20-Jun-1989 #text\_change 15-Oct-1999  
C:Accession: B28474  
R:Freid, J.F.; Hengstenberg, W.; Finkeldei, U.; Stewart, G.C.  
FEBS Lett. 258, 16444-16449, 1989  
A:Title: Identification of the genes for the lactose-specific components of the phosphotransferase system enzyme II (EC 2.7.1.69), lactose-specific, factor II - Step

A: Reference number: A92619; MUID:88059023  
 A: Accession: B28474  
 A: Molecule type: DNA  
 A: Species: *Prunus serotina* (black cherry)  
 C: Date: 21-Jul-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C: Accession: T48795  
 A: Cross-references: GB:J03479; NID:g153036; PIDN:AA26649.1; PID:g153038  
 C: Keywords: phosphotransferase

Query Match 100.0%; Score 33; DB 2; Length 572;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGTSG 6  
 Db 478 GCGTSG 483

RESULT 11  
 T10766  
 A: Reference number: EC 4.1.2.10  
 A: Accession: A00000 (imported) - black cherry  
 C: Species: *Prunus serotina* (black cherry)  
 C: Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
 C: Accession: T50766  
 R: Hu, Z.; Poulton, J. E.  
 A: Molecule type: DNA  
 A: Species: *Prunus serotina* (black cherry)  
 A: Accession: T50766  
 A: Cross-references: Translated from GB/EMBL/DBJ  
 A: Status: Preliminary  
 A: Molecule type: DNA  
 A: Residues: 1-574 <NUZ>  
 A: Cross-references: EMBL:AF043186; PIDN:AA02265.1  
 C: Superfamily: alcohol oxidase  
 C: Keywords: aldehyde lyase; carbon-carbon lyase

Query Match 100.0%; Score 33; DB 2; Length 574;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGTSG 6  
 Db 61 GCGTSG 66

RESULT 12  
 B32657  
 A: Reference number: EC 2.7.1.69  
 A: Accession: B23697  
 C: Species: *Lactobacillus casei*  
 C: Date: 04-Oct-1991 #sequence\_revision 04-Oct-1991 #text\_change 15-Oct-1999  
 C: Accession: B23697  
 R: Alpert, C.A.; Chaasy, B.M.  
 J Biol. Chem. 265, 22561-22568, 1990  
 A: Molecule type: DNA  
 A: Species: *Lactobacillus casei*  
 A: Accession: B23697  
 A: Cross-references: MUID:91093108  
 C: Keywords: phosphotransferase; transmembrane protein

Query Match 100.0%; Score 33; DB 2; Length 577;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGTSG 6  
 Db 485 GCGTSG 490

RESULT 13

T48795  
 A: Reference number: A92619; MUID:88059023  
 A: Accession: B28474  
 A: Molecule type: DNA  
 A: Species: *Prunus serotina* (black cherry)  
 C: Date: 21-Jul-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C: Accession: T48795  
 A: Cross-references: GB:J03479; NID:g153036; PIDN:AA26649.1; PID:g153038  
 C: Keywords: phosphotransferase

Query Match 100.0%; Score 33; DB 2; Length 578;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGTSG 6  
 Db 521 GCGTSG 526

RESULT 14  
 F84855  
 A: Reference number: A92619; MUID:88059023  
 A: Accession: B28474  
 A: Molecule type: DNA  
 A: Species: *Prunus serotina* (black cherry)  
 C: Date: 21-Jul-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C: Accession: T48795  
 A: Cross-references: GB:J03479; NID:g153036; PIDN:AA26649.1; PID:g153038  
 C: Keywords: phosphotransferase

Query Match 100.0%; Score 33; DB 2; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.8e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCGTSG 6  
 Db 203 GCGTSG 208

RESULT 15  
 S52485  
 A: Reference number: A92619; MUID:88059023  
 A: Accession: B28474  
 A: Molecule type: DNA  
 A: Species: *Prunus serotina* (black cherry)  
 C: Date: 21-Jul-2000 #sequence\_revision 05-May-2000 #text\_change 05-May-2000  
 C: Accession: T48795  
 A: Cross-references: GB:J03479; NID:g153036; PIDN:AA26649.1; PID:g153038  
 C: Keywords: phosphotransferase

A:Residues: 1-625 <BON>  
A:Cross-references: EMBL:Z48475; NID:q68357  
C:Gene: GDB:GCCR  
A:Gene: GDB:GCCR  
A:Cross-references: GDB:207312; OMIM:600842  
A:Map position: 2p23.3-2p23.2  
C:Superfamily: glucokinase regulator

Query Match	100.0%;	Score 33;	DB 1;	Length 625;
Best Local Similarity	100.0%;	Pred. NO.	1	9e+02.

Qy	1	GGGTSG	6
Db	106	GGGTSG	111

RESULT 16  
S4175 glucocortisol receptor 1 - rat  
S4175 glucocortisol receptor 1 - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #date  
C:Accession: S4175  
R:Belchew, M.; VandeKerkhove, J.; van Schaftingen, E.  
R:Slett, 339, 312-315, 1994

A:Reference number: SI1745; MIMD:94150054  
A:Accession: SI1745  
A:Molecule type: mRNA  
A:Residues: 1-677 <DET>  
A:Superfamily: glucocorticoid inducible protein

Qy	1	GGTSG	6
		11111	
Db	106	GGTSG	111

R29266 17  
 keratin, 65k type II cytoskeletal - human  
 C.Species: Homo sapiens (man)  
 C.Date: 31-Dec-1988 #sequence\_revision 30-Sep-1991 #text\_change 20-Jun-2000  
 C.Accession: A29666  
 J.Lindegren, R.M.: Sylvestre, Y.R.: Friedberg, I.M.: Blumenberg, M.  
 A.Title: Evolution of the 30-32k, 1997  
 A.Reference numbers: A29666, M01D018754235  
 A.Accession: A29666  
 A: Molecule type: DNA  
 A: Residues: 1-629 <KLI>  
 A29694.1: PID:NI05418: NID:934040: PIDN:CA28991.1: PID:G34041: GB:X05419: NID:9  
 A29694.1: PID:G135193: NID:934043: PIDN:CA28995.1: PID:G135194: GB:X05421: NID:934047: NID:934048: NID:934049: NID:934050: NID:934051: NID:934052: NID:934053: NID:934054: NID:934055: NID:934056: NID:934057: NID:934058: NID:934059: NID:934060: NID:934061: NID:934062: NID:934063: NID:934064: NID:934065: NID:934066: NID:934067: NID:934068: NID:934069: NID:934070: NID:934071: NID:934072: NID:934073: NID:934074: NID:934075: NID:934076: NID:934077: NID:934078: NID:934079: NID:934080: NID:934081: NID:934082: NID:934083: NID:934084: NID:934085: NID:934086: NID:934087: NID:934088: NID:934089: NID:934090: NID:934091: NID:934092: NID:934093: NID:934094: NID:934095: NID:934096: NID:934097: NID:934098: NID:934099: NID:934100: NID:934101: NID:934102: NID:934103: NID:934104: NID:934105: NID:934106: NID:934107: NID:934108: NID:934109: NID:934110: NID:934111: NID:934112: NID:934113: NID:934114: NID:934115: NID:934116: NID:934117: NID:934118: NID:934119: NID:934120: NID:934121: NID:934122: NID:934123: NID:934124: NID:934125: NID:934126: NID:934127: NID:934128: NID:934129: NID:934130: NID:934131: NID:934132: NID:934133: NID:934134: NID:934135: NID:934136: NID:934137: NID:934138: NID:934139: NID:934140: NID:934141: NID:934142: NID:934143: NID:934144: NID:934145: NID:934146: NID:934147: NID:934148: NID:934149: NID:934150: NID:934151: NID:934152: NID:934153: NID:934154: NID:934155: NID:934156: NID:934157: NID:934158: NID:934159: NID:934160: NID:934161: NID:934162: NID:934163: NID:934164: NID:934165: NID:934166: NID:934167: NID:934168: NID:934169: NID:934170: NID:934171: NID:934172: NID:934173: NID:934174: NID:934175: NID:934176: NID:934177: NID:934178: NID:934179: NID:934180: NID:934181: NID:934182: NID:934183: NID:934184: NID:934185: NID:934186: NID:934187: NID:934188: NID:934189: NID:934190: NID:934191: NID:934192: NID:934193: NID:934194: NID:934195: NID:934196: NID:934197: NID:934198: NID:934199: NID:934200: NID:934201: NID:934202: NID:934203: NID:934204: NID:934205: NID:934206: NID:934207: NID:934208: NID:934209: NID:934210: NID:934211: NID:934212: NID:934213: NID:934214: NID:934215: NID:934216: NID:934217: NID:934218: NID:934219: NID:934220: NID:934221: NID:934222: NID:934223: NID:934224: NID:934225: NID:934226: NID:934227: NID:934228: NID:934229: NID:934230: NID:934231: NID:934232: NID:934233: NID:934234: NID:934235: NID:934236: NID:934237: NID:934238: NID:934239: NID:934240: NID:934241: NID:934242: NID:934243: NID:934244: NID:934245: NID:934246: NID:934247: NID:934248: NID:934249: NID:934250: NID:934251: NID:934252: NID:934253: NID:934254: NID:934255: NID:934256: NID:934257: NID:934258: NID:934259: NID:934260: NID:934261: NID:934262: NID:934263: NID:934264: NID:934265: NID:934266: NID:934267: NID:934268: NID:934269: NID:934270: NID:934271: NID:934272: NID:934273: NID:934274: NID:934275: NID:934276: NID:934277: NID:934278: NID:934279: NID:934280: NID:934281: NID:934282: NID:934283: NID:934284: NID:934285: NID:934286: NID:934287: NID:934288: NID:934289: NID:934290: NID:934291: NID:934292: NID:934293: NID:934294: NID:934295: NID:934296: NID:934297: NID:934298: NID:934299: NID:934300: NID:934301: NID:934302: NID:934303: NID:934304: NID:934305: NID:934306: NID:934307: NID:934308: NID:934309: NID:934310: NID:934311: NID:934312: NID:934313: NID:934314: NID:934315: NID:934316: NID:934317: NID:934318: NID:934319: NID:934320: NID:934321: NID:934322: NID:934323: NID:934324: NID:934325: NID:934326: NID:934327: NID:934328: NID:934329: NID:934330: NID:934331: NID:934332: NID:934333: NID:934334: NID:934335: NID:934336: NID:934337: NID:934338: NID:934339: NID:934340: NID:934341: NID:934342: NID:934343: NID:934344: NID:934345: NID:934346: NID:934347: NID:934348: NID:934349: NID:934350: NID:934351: NID:934352: NID:934353: NID:934354: NID:934355: NID:934356: NID:934357: NID:934358: NID:934359: NID:934360: NID:934361: NID:934362: NID:934363: NID:934364: NID:934365: NID:934366: NID:934367: NID:934368: NID:934369: NID:934370: NID:934371: NID:934372: NID:934373: NID:934374: NID:934375: NID:934376: NID:934377: NID:934378: NID:934379: NID:934380: NID:934381: NID:934382: NID:934383: NID:934384: NID:934385: NID:934386: NID:934387: NID:934388: NID:934389: NID:934390: NID:934391: NID:934392: NID:934393: NID:934394: NID:934395: NID:934396: NID:934397: NID:934398: NID:934399: NID:934400: NID:934401: NID:934402: NID:934403: NID:934404: NID:934405: NID:934406: NID:934407: NID:934408: NID:934409: NID:934410: NID:934411: NID:934412: NID:934413: NID:934414: NID:934415: NID:934416: NID:934417: NID:934418: NID:934419: NID:934420

Query Match	100.0%	Score 33	DB 2	Length 629
Best Local Similarity	100.0%	Prod. No. 1.9e+02		
Matches 6	Conservative 0	Mismatches 0	Indels 0	Gaps 0
0y	1 GGCTSG 6			

QY 1 GGGTSG 6  
111111

Db 578 GGGTSG 583

RESULT 18  
Q0BEM1  
U52 protein - human cytomegalovirus (strain AD169)  
N: Alternate names: HFRF1 protein  
C: Species: human cytomegalovirus

Query Match	100.0%;	Score 33;	DB 1;	Length 668;
Best Local Similarity	100.0%;	Pred. No. 2c+02;		
Matches	6;	Conservative	0;	Mismatches 0; Indels 0; Gaps 0;

QY	1	GGGTSG	6
Db	422	GGGTSG	427

RESULT 19  
T27941  
hypothetical protein ZK662.3 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 18-Feb-2000  
R:Accession: T27941  
3 submitted to the EMBL Data Library, August 1996  
A:Reference number: Z20444  
A:Accession: T27941  
A:Status: preliminary: translated from BR/BNL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-780 <MIL>  
A:Cross-references: EMBL:Z79604; PIDN:CA01900.1; GSPDB:GN00028; CSDP:ZK662.3  
A:Experimental source: clone ZK662  
C:Genetics:  
A:Gene: CSDP:ZK662.3  
A:Map position: X  
A:Features: 502/1; 56/1; 98/3; 176/3; 206/3; 289/2; 311/1; 341/3; 444/3; 507/3; 655/1;

```

Query Match      100.0%; Score 33; DB 2; Length 780;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY	1	GGGTSC	6
Db	64	GGGTSC	69

**RESULT** 20  
 B35905  
 endopeptidase Clp (PC 3.4.21.-) ATP-binding chain cdk9, chloroplast [similarity] - **cdk9**  
 N:Alternate names: ATP-dependent Clp proteinase regulatory chain B; Cdk9 protein  
 N:Contents: Adenosine triphosphatase (EC 3.6.1.3)  
 C:Species: *Adenosiraon eschscholmii* (tomato)  
 C:Date: 30-Jun-1993 #sequence\_revision 30-Jun-1993 #texl\_change 19-jun-2001  
 C:Accession: B35905

R:Gottesman, S.; Squires, C.; Plocherky, E.; Cartington, M.; Hobbs, M.; Matlick, J.S.; D  
Proc. Natl. Acad. Sci. U.S.A. 87, 3513-3517, 1990  
A:Title: Conservation of the regulatory subunit for the Clp ATP-dependent protease in pr  
A:Accession: AF039051; MIM:90259044  
A:Residues: 1-923 <CONT>  
A:Molecule type: DNA  
A:Cross-references: GB:M32604; NID:g170432; PIRN:MAA3161.1; PIR:g170435  
C:Genetics:  
A:Gene: cddA  
A:Map position: 12  
C:Function:  
A:Description: allows clp to hydrolyze polypeptides and proteins, probably by a chaper  
A:activity: ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
C:superfamily: chaperonin; molecular chaperone; nucleotide binding; P-loop  
C:Keywords: ATP; chaperonin; hydrolysis; molecular chaperone; nucleotide binding; P-loop  
F:300-307/Region: nucleotide-binding motif A (P-loop)  
F:367-372/Region: nucleotide-binding motif B (P-loop)  
F:411-415/Region: nucleotide-binding motif A (P-loop)  
F:711-715/Region: nucleotide-binding motif B (P-loop)  
F:306/Binding site: ATP (lys) #status predicted  
F:645/Binding site: ATP (lys) #status predicted

Query Match 100.0%; Score 33; DB 1; Length 923;  
Best Local Similarity 100.0%; Pred. No. 2, 7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 244 GCGTSG 249

RESULT 21  
A:Accession: AF039051  
A:Title: Conservation of the regulatory subunit for the Clp ATP-dependent protease in pr  
A:Accession: AF039051; MIM:90259044  
A:Residues: 1-923 <CONT>  
A:Molecule type: DNA  
A:Cross-references: GB:M32603; NID:g170432; PIRN:MAA3160.1; PIR:g170433  
C:Genetics:  
A:Gene: cddA  
A:Map position: 3  
C:Function:  
A:Description: allows clp to hydrolyze polypeptides and proteins, probably by a chaper  
A:activity: ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
C:superfamily: chaperonin; molecular chaperone; nucleotide binding; P-loop  
C:Keywords: ATP; chaperonin; hydrolysis; molecular chaperone; nucleotide binding; P-loop  
F:300-307/Region: nucleotide-binding motif A (P-loop)  
F:367-372/Region: nucleotide-binding motif B (P-loop)  
F:411-415/Region: nucleotide-binding motif A (P-loop)  
F:711-715/Region: nucleotide-binding motif B (P-loop)  
F:306/Binding site: ATP (lys) #status predicted  
F:645/Binding site: ATP (lys) #status predicted

Query Match 100.0%; Score 33; DB 1; Length 926;  
Best Local Similarity 100.0%; Pred. No. 2, 7e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 246 GCGTSG 251

RESULT 22  
A:Accession: AF039051  
A:Title: Conservation of the regulatory subunit for the Clp ATP-dependent protease in pr  
A:Accession: AF039051; MIM:90259044  
A:Residues: 1-923 <CONT>  
A:Molecule type: DNA  
A:Cross-references: GB:M32604; NID:g170432; PIRN:MAA3161.1; PIR:g170435  
C:Genetics:  
A:Gene: cddA  
A:Map position: 12  
C:Function:  
A:Description: allows clp to hydrolyze polypeptides and proteins, probably by a chaper  
A:activity: ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
C:superfamily: chaperonin; molecular chaperone; nucleotide binding; P-loop  
C:Keywords: ATP; chaperonin; hydrolysis; molecular chaperone; nucleotide binding; P-loop  
F:300-307/Region: nucleotide-binding motif A (P-loop)  
F:367-372/Region: nucleotide-binding motif B (P-loop)  
F:411-415/Region: nucleotide-binding motif A (P-loop)  
F:711-715/Region: nucleotide-binding motif B (P-loop)  
F:306/Binding site: ATP (lys) #status predicted  
F:645/Binding site: ATP (lys) #status predicted

Query Match 100.0%; Score 33; DB 2; Length 1192;  
Best Local Similarity 100.0%; Pred. No. 3, 4e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCGTSG 6  
DB 101 GCGTSG 106

RESULT 23  
A:Accession: AF039051  
A:Title: Conservation of the regulatory subunit for the Clp ATP-dependent protease in pr  
A:Accession: AF039051; MIM:90259044  
A:Residues: 1-923 <CONT>  
A:Molecule type: DNA  
A:Cross-references: GB:M32604; NID:g170432; PIRN:MAA3161.1; PIR:g170435  
C:Genetics:  
A:Gene: cddA  
A:Map position: 3  
C:Function:  
A:Description: allows clp to hydrolyze polypeptides and proteins, probably by a chaper  
A:activity: ATP hydrolysis is required for Clp hydrolysis of proteins but not of smaller  
C:superfamily: chaperonin; molecular chaperone; nucleotide binding; P-loop  
C:Keywords: ATP; chaperonin; hydrolysis; molecular chaperone; nucleotide binding; P-loop  
F:300-307/Region: nucleotide-binding motif A (P-loop)  
F:367-372/Region: nucleotide-binding motif B (P-loop)  
F:411-415/Region: nucleotide-binding motif A (P-loop)  
F:711-715/Region: nucleotide-binding motif B (P-loop)  
F:306/Binding site: ATP (lys) #status predicted  
F:645/Binding site: ATP (lys) #status predicted



iller, L.; Groebeck, E.J.; Davis, N.W.; Lim, A.; Dimantanta, E.; Potamomusis, K.; Apod  
Nature 409, 529-533, 2001  
A>Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: AB5480; M01D:21074935; PMID:11206551  
A:Accession: D85915  
A>Status: Preliminary  
A:Source: *Escherichia coli* O157:H7  
A:Cross-references: GR:AE005174; NID:91251713; PIDN:ANG57784.1; GSPDB:GN0145; UNCP  
A:Experimental source: strain O157:H7, substrain EDJ933  
C:Genetics:  
A:Gene: nrdI  
C:Superfamily: *Bacillus subtilis* conserved hypothetical protein yosM

Query Match 90.9%; Score 30; DB 2; Length 136;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTGC 6  
|||||  
Db 51 GGCTTC 56

RESULT 26  
T46122 305 ribosomal protein S9-1like - *Arabidopsis thaliana*  
N:Alternate names: protein T2J13.80  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
A>Title: Sequence revision 04-Feb-2000 \*text\_change 18-Feb-2000  
R:Release: M.; Gabli, C.; Mueller-Auer, S.; Schaefer, M.; Zipp, M.; Mewes, H.W.; Mayer  
submitted to the Protein Sequence Database, November 1999  
A:Reference number: Z23023  
A:Accession: T46122  
A>Status: Preliminary  
A:Molecule type: NRD  
A:Length: 305 residues; 114 kDa  
A:Cross-references: EMBL:AL132967  
C:Genetics:  
A:Map position: 3  
A:Introns: 105/1; 114/1  
A:Note: T2J13.80  
C:Superfamily: *Escherichia coli* ribosomal protein S9

Query Match 90.9%; Score 30; DB 2; Length 143;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGCTGC 6  
|||||  
Db 82 GGCTTC 87

RESULT 27  
F84182 hypothetical protein Yng0219h [Imported] - *Halobacterium* sp. NRC-1  
C:Species: *Halobacterium* sp. NRC-1  
A>Title: Sequence revision 02-Feb-2001 \*text\_change 02-Feb-2001  
C:Date: 02-Feb-2001  
R:Release: M.V.; Kennedy, S.P.; Mohitias, G.G.; Benquist, B.; Pan, M.; Shultz, H.D.; Lasty  
R:Release: M.V.; Kennedy, S.P.; Mohitias, G.G.; Benquist, B.; Pan, M.; Shultz, H.D.; Lasty  
L:Release: M.V.; Kennedy, S.P.; Mohitias, G.G.; Benquist, B.; Pan, M.; Shultz, H.D.; Lasty  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Demis, P.P.; Omer, A.D.; Ehardt, E.; Lowe, T.N.  
A>Title: Genome sequence of *Halobacterium* species NRC-1.  
A:Accession: F84182  
A:Experimental source: strain 84180; M01D:20504405  
A>Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-185 <STO>

A:Cross-references: GB:AE004437; NID:g10579865; PIDN:AG18826.1; GSPDB:GN00138  
 C:Genetics:  
 A:Gene: VNC02198

Query Match 90.9% Score 30; DB 2; Length 185;  
 Best Local Similarity 83.3% Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSC 6  
 22 GGGTGA 27

RESULT 28

Conserved hypothetical protein YJL055W - yeast (Saccharomyces cerevisiae)  
 M:Accession names: hypothetical protein J1148  
 C:Date: 08-20-1999 #sequence\_revision 08-Sep-1995 #text\_change 08-Oct-1999  
 C:Accession: S56827  
 R:Publ. T.M.: Aljinovic, G.  
 Submitted to the Protein Sequence Database, September 1995  
 A:Reference number: S56793  
 A:Accession: S56827  
 A:Molecule type: DNA  
 A:Status: preliminary  
 A:Cross-references: EMBL:Z49330; NID:g1008194; PIDN:CA89346.1; PID:g1008195; MIPS:YJL055W  
 C:Genetics:  
 A:Map position: 10L  
 C:Superfamily: yeast conserved hypothetical protein YJL055W

Query Match 90.9% Score 30; DB 2; Length 245;  
 Best Local Similarity 83.3% Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSC 6  
 56 GGGTGA 61

RESULT 29

glycine-rich protein grp3 - maize  
 C:Species: Zea mays (maize)  
 C:Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 21-Jul-2000  
 C:Accession: T03371  
 R:Publ. T.M.: Wolff, D.; Felt, G.  
 Plant Mol. Biol. 35: 799-802, 1998  
 A:Title: Root-specific expression of a Zea mays gene encoding a novel glycine-rich protein  
 A:Reference number: Z14904; MUID:98187261  
 A:Accession: T03371  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: mRNA  
 A:Status: preliminary  
 A:Cross-references: EMBL:Y07781; NID:g1552070; PIDN:CA69104.1; PID:g1532071  
 A:Experimental source: strain DK05; root  
 C:Genetics:  
 A:Gene: grp3  
 C:Superfamily: glycine-rich cell wall structural protein 1

Query Match 90.9% Score 30; DB 2; Length 256;  
 Best Local Similarity 83.3% Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSC 6  
 92 GGGTGA 97

RESULT 30

F75300  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: F75300  
 R:Miller, O.; Eisen, J.A.; Haldberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.  
 Science 286: 1571-1577, 1999  
 S: Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 A:Title: Genome sequence of the radiolabel-resistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036996  
 A:Accession: F75300  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1276 <WHI>  
 A:Cross-references: GB:AE002055; GB:AE000513; NID:g6460024; PIDN:AMF11780.1; PID:g646  
 C:Genetics:  
 A:Gene: DR2232  
 A:Map position: 1

Query Match 90.9% Score 30; DB 2; Length 276;  
 Best Local Similarity 83.3% Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSC 6  
 61 GGGTGA 66

RESULT 31

hypothetical protein eutY (imported) - Escherichia coli (strain 0157:H7)  
 G85887  
 C:Species: Escherichia coli  
 C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 31-Mar-2001  
 C:Accession: G85887  
 R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
 11er, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Diallanita, E.; Potamousis, K.; Apoda  
 Nature 409: 529-533, 2001  
 A:Title: Genome sequence of enterohemorrhagic Escherichia coli 0157:H7.  
 A:Reference number: A85480; MUID:21074935; PMID:11206551  
 A:Accession: G85887  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1278 <STO>  
 A:Cross-references: GB:AE005174; NID:g12516826; PIDN:AG57563.1; GSPDB:GN00145; UNCP:  
 A:Experimental source: strain 0157:H7, substrain EDL933  
 C:Genetics:  
 A:Gene: eutY

Query Match 90.9% Score 30; DB 2; Length 278;  
 Best Local Similarity 83.3% Pred. No. 2.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGTSC 6  
 145 GGGTGA 150

RESULT 32

E65020  
 ethanolamine utilization protein EutY - Escherichia coli (strain K-12)  
 C:Species: Escherichia coli  
 C:Date: 17-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
 C:Accession: E65020  
 R:Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
 Science 277: 1453-1462, 1997  
 A:Title: The complete genome sequence of Escherichia coli K-12.  
 A:Reference number: A64720; MUID:97426617  
 A:Accession: E65020

A:Status: Preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Accession: U000003.2; GB:U00006; NID:q1286789; PID:NAC75507.1; PID:q1788796;  
 A:Cross-references: GB:AF000332;  
 A:Experimental source: strain K-12, substrain MG1655  
 C:Genes: elc1

Query Match  
 Best Local Similarity 90.9%; Score 30; DB 2; Length 278;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGTSG 6  
 |||||:1  
 145 GCGTTG 150

RESULT 33

JC6502  
 peccate lyase (EC 4.2.2.2) - Amycolata sp.

C:Species: Amycolata sp.

C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: JC6502 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

C:Accession: JC6502 #sequence\_revision 16-Jul-1999 #text\_change 21-Jul-2000

A:Title: Cloning, sequence and expression of the pel gene from an Amycolata sp.

A:Reference number: JC6502; MUID:98087416

A:Accession: JC6502

A:Status: Preliminary

A:Molecule type: DNA

A:Accession: JC6502

A:Cross-references: GB:AF000332; NID:92196706; PID:NAC38059.1; PID:q2196707

C:Keywords: carbon-oxygen lyase

Query Match  
 Best Local Similarity 90.9%; Score 30; DB 2; Length 310;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGTSG 6  
 |||||:1  
 38 GCGTTG 43

RESULT 34

hypothetical protein fl3A7.1 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999

C:Accession: T20807

R:McMurray, A.  
 submitted to the EMBL Data Library, March 1997

A:Accession: T20807

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-340 <MID>

A:Cross-references: EMBL:293377; PID:NAC80573.1; GSPDB:GN00023; CESP:FL3A7.1

A:Experimental source: clone FL3A7

A:Gene: CESP:FL3A7.1

A:Map position: 5

A:introns: 6/2; 44/2; 80/3

Query Match  
 Best Local Similarity 90.9%; Score 30; DB 2; Length 340;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGTSG 6  
 |||||:1  
 224 GCGTAG 229

RESULT 35  
 T34954  
 Probable UDP-N-acetylglucosamine-N-acetylmutamyl-(pentapeptide) pyrophosphoryl-undec

C:Species: Streptomyces coelicolor

C:Date: 01-Jun-1999 #sequence\_revision 05-Nov-1999 #text\_change 21-Jan-2000

C:Accession: T34954

R:Saunders, D.C.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M

submitted to the EMBL Data Library, August, 1999

A:Reference number: 221563

A:Accession: T34954

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-364 <MID>

A:Cross-references: EMBL:AL109653; PID:CAH51993.1; GSPDB:GN00070; SCOPDB:SC4A10.17c

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: mutG; SCOPDB:SC4A10.17c

C:Superfamily: mutG protein

Query Match  
 Best Local Similarity 90.9%; Score 30; DB 2; Length 364;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGTSG 6  
 |||||:1  
 7 GCGTAG 12

RESULT 36

I57555

C-Mat protein - mouse

C:Species: Mus sp. (mouse)

C:Date: 29-May-1998 #sequence\_revision 29-May-1998 #text\_change 16-Jul-1999

C:Accession: I57555

R:Kirschner, C.; Morgan, J.I.

Mol. Cell. Biol. 15: 246-254, 1995

A:Title: The mat proto-oncogene stimulates transcription from multiple sites in a pro

A:Reference number: I57555; MUID:95097997

A:Accession: I57555

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-370 <RES>

A:Cross-references: GB:SV4567; NID:9807054; PID:AB32820.1; PID:9807055

C:Genetics:

A:Gene: c-maf

C:Superfamily: maf transforming protein; maf homology

F:259-348/Domain: maf homology <MAF>

Query Match  
 Best Local Similarity 90.9%; Score 30; DB 2; Length 370;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Db 1 GCGTSG 6  
 |||||:1  
 233 GCGTAG 238

RESULT 37

B34261

alanine dehydrogenase (EC 1.4.1.1) - Bacillus stearothermophilus

C:Species: Bacillus stearothermophilus

C:Date: 01-Jun-1990 #sequence\_revision 01-Jun-1990 #text\_change 11-Jun-1999

C:Accession: B34261

R:Wada, S.; Tanizawa, Y.; Tanaka, H.; Soda, K.

Biochimically 29, 1009-1015, 1990

A:Title: Alanine dehydrogenases from two Bacillus species with distinct thermostabili

ty and dehydrogenase

A:Reference number: A34261; MUID:90254112

A:Accession: B34261

A:status: Preliminary  
 A:molecule type: DNA  
 A:Residues: 1-372 <RND>  
 A:Cross-references: GB:R33239; NID:9142454; PIRN:AAA2211.1; PIR:9142455  
 C:Species: Mesocricetus auratus (golden hamster)  
 C>Date: 21-Sep-1993 #sequence,revision 18-Nov-1994 #text,change 16-Jul-1999  
 C:Accession: B46233  
 R:Gorman, M.S.; Wang, J.; Chaudhry, R.B.; Rutter, W.J.  
 A:Title: Spontaneous activation of the insulin gene by a LIM-homeo domain protein and a  
 A:Reference number: A46233; M0ID:93051335  
 A:Accession: B46233  
 A:status: Preliminary  
 A:molecule type: mRNA  
 A:Residues: 1-382 <GER>  
 A:Cross-references: GB:R81406; NID:9587460; PIRN:CAA57163.1; PIR:9587461  
 C:Species: Mus musculus (house mouse)  
 C>Date: 14-Jul-2000 #sequence,revision 20-Apr-2000 #text,change 20-Apr-2000  
 C:Accession: T48109  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemke, K.  
 A:Title: The LIM metal-binding repeat homology <LIM>  
 A:Reference number: A48233; M0ID:93051335  
 A:Accession: B48233  
 A:status: Preliminary  
 A:molecule type: DNA  
 A:Residues: 1-382 <RND>  
 A:Cross-references: EMBL:AL38648  
 A:Experimental source: Cultivar Columbia; BAC clone F16M2  
 C:Accession: F170-198/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 90.9%; Score 30; DB 2; Length 372;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GGGTSG 6  
 DB 175 GGGTAC 180

RESULT 38

Transcription factor Lmx-1 - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C>Date: 21-Sep-1993 #sequence,revision 18-Nov-1994 #text,change 16-Jul-1999  
 C:Accession: B46233  
 R:Gorman, M.S.; Wang, J.; Chaudhry, R.B.; Rutter, W.J.  
 A:Title: Spontaneous activation of the insulin gene by a LIM-homeo domain protein and a  
 A:Reference number: A46233; M0ID:93051335  
 A:Accession: B46233  
 A:status: Preliminary  
 A:molecule type: mRNA  
 A:Residues: 1-382 <GER>  
 A:Cross-references: GB:R81406; NID:9587460; PIRN:CAA57163.1; PIR:9587461  
 C:Species: Mus musculus (house mouse)  
 C>Date: 14-Jul-2000 #sequence,revision 20-Apr-2000 #text,change 20-Apr-2000  
 C:Accession: T48109  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemke, K.  
 A:Title: The LIM metal-binding repeat homology <LIM>  
 A:Reference number: A48233; M0ID:93051335  
 A:Accession: B48233  
 A:status: Preliminary  
 A:molecule type: DNA  
 A:Residues: 1-382 <RND>  
 A:Cross-references: EMBL:AL38648  
 A:Experimental source: Cultivar Columbia; BAC clone F16M2  
 C:Accession: F170-198/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 90.9%; Score 30; DB 1; Length 382;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GGGTSG 6  
 DB 277 GGGTAC 282

RESULT 39

Transcription factor Lmx-1 - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C>Date: 21-Sep-1993 #sequence,revision 18-Nov-1994 #text,change 16-Jul-1999  
 C:Accession: B46233  
 R:Gorman, M.S.; Wang, J.; Chaudhry, R.B.; Rutter, W.J.  
 A:Title: Spontaneous activation of the insulin gene by a LIM-homeo domain protein and a  
 A:Reference number: A46233; M0ID:93051335  
 A:Accession: B46233  
 A:status: Preliminary  
 A:molecule type: mRNA  
 A:Residues: 1-382 <GER>  
 A:Cross-references: GB:R81406; NID:9587460; PIRN:CAA57163.1; PIR:9587461  
 C:Species: Mus musculus (house mouse)  
 C>Date: 14-Jul-2000 #sequence,revision 20-Apr-2000 #text,change 20-Apr-2000  
 C:Accession: T48109  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemke, K.  
 A:Title: The LIM metal-binding repeat homology <LIM>  
 A:Reference number: A48233; M0ID:93051335  
 A:Accession: B48233  
 A:status: Preliminary  
 A:molecule type: DNA  
 A:Residues: 1-382 <RND>  
 A:Cross-references: EMBL:AL38648  
 A:Experimental source: Cultivar Columbia; BAC clone F16M2  
 C:Accession: F170-198/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 90.9%; Score 30; DB 2; Length 382;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GGGTSG 6  
 DB 35 GGGTAC 40

RESULT 40

Transcription factor Lmx-1 - golden hamster  
 C:Species: Mesocricetus auratus (golden hamster)  
 C>Date: 21-Sep-1993 #sequence,revision 18-Nov-1994 #text,change 16-Jul-1999  
 C:Accession: B46233  
 R:Gorman, M.S.; Wang, J.; Chaudhry, R.B.; Rutter, W.J.  
 A:Title: Spontaneous activation of the insulin gene by a LIM-homeo domain protein and a  
 A:Reference number: A46233; M0ID:93051335  
 A:Accession: B46233  
 A:status: Preliminary  
 A:molecule type: mRNA  
 A:Residues: 1-382 <GER>  
 A:Cross-references: GB:R81406; NID:9587460; PIRN:CAA57163.1; PIR:9587461  
 C:Species: Mus musculus (house mouse)  
 C>Date: 14-Jul-2000 #sequence,revision 20-Apr-2000 #text,change 20-Apr-2000  
 C:Accession: T48109  
 R:Rieger, M.; Mueller-Auer, S.; Zipp, M.; Schaefer, M.; Mewes, H.W.; Rudd, S.; Lemke, K.  
 A:Title: The LIM metal-binding repeat homology <LIM>  
 A:Reference number: A48233; M0ID:93051335  
 A:Accession: B48233  
 A:status: Preliminary  
 A:molecule type: DNA  
 A:Residues: 1-382 <RND>  
 A:Cross-references: EMBL:AL38648  
 A:Experimental source: Cultivar Columbia; BAC clone F16M2  
 C:Accession: F170-198/Region: beta-alpha-beta NAD nucleotide-binding fold

Query Match 90.9%; Score 30; DB 2; Length 384;  
 Best Local Similarity 83.3%; Pred. No. 3.8e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

1 GGGTSG 6  
 DB 64 GGGTAC 69

Search completed: February 4, 2002, 08:00:52  
 Job time: 36 sec



**THIS PAGE BLANK (USPTO)**